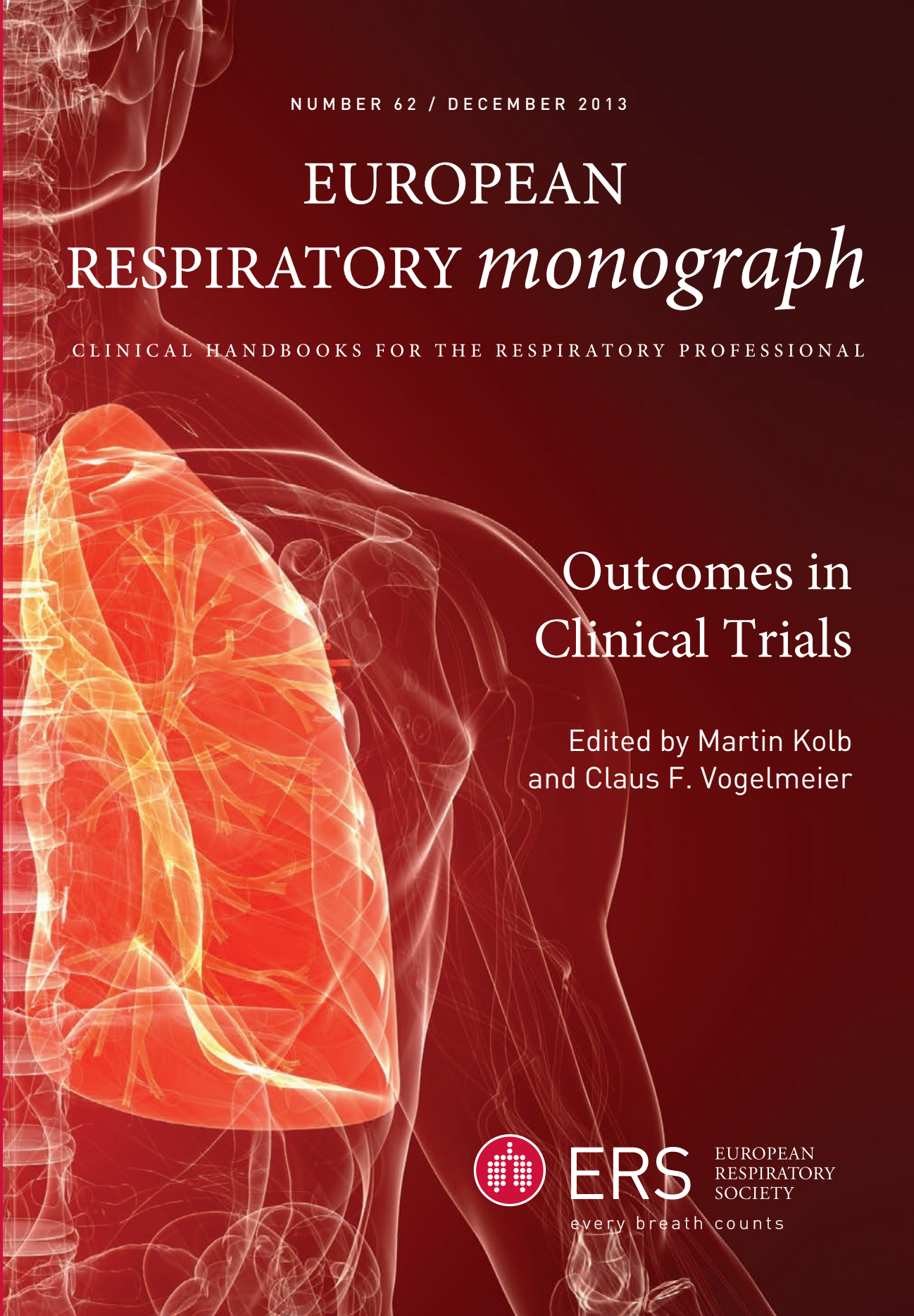
NUMBER 62 / DECEMBER 2013

EUROPEAN RESPIRATORY *monograph*

CLINICAL HANDBOOKS FOR THE RESPIRATORY PROFESSIONAL



Outcomes in Clinical Trials

Edited by Martin Kolb and Claus F. Vogelmeier

ERS EUROPEAN RESPIRATORY SOCIETY
every breath counts

European Respiratory Monograph 62, December 2013

Outcomes in Clinical Trials

Edited by
Martin Kolb and Claus F. Vogelmeier

Editor in Chief
Tobias Welte

Published by European Respiratory Society ©2013
December 2013
Print ISBN: 978-1-84984-044-6
Online ISBN: 978-1-84984-045-3
Print ISSN: 1025-448x
Online ISSN: 2075-6674
Printed by Page Bros Ltd, Norwich, UK

Managing Editors: Rachel White and Catherine Pumphrey
European Respiratory Society
442 Glossop Road, Sheffield,
S10 2PX, UK
Tel: 44 114 2672860
E-mail: Monograph@ersj.org.uk

All material is copyright to European Respiratory Society. It may not be reproduced in any way including electronic means without the express permission of the company.

Statements in the volume reflect the views of the authors, and not necessarily those of the European Respiratory Society, editors or publishers.

THOMSON REUTERS BOOK CITATION INDEX INDEXED

FSC www.fsc.org MIX Paper from responsible sources FSC® C023114

This book is one in a series of *European Respiratory Monographs*. Each individual issue provides a comprehensive overview of one specific clinical area of respiratory health, communicating information about the most advanced techniques and systems required for its investigation. It provides factual and useful scientific detail, drawing on specific case studies and looking into the diagnosis and management of individual patients. Previously published titles in this series are listed at the back of this *Monograph*.

Contents

Number 62 December 2013

Preface v

Guest Editors vii

Introduction ix

Outcomes in clinical trials: specific disease areas

1. Chronic obstructive pulmonary disease 1
Emiel F.M. Wouters

2. Asthma 19
Julia Koopmans and Christian Taube

3. Idiopathic pulmonary fibrosis 37
Toby M. Maher

4. Cystic fibrosis 54
Reshma Amin and Felix Ratjen

5. Pulmonary arterial hypertension 70
Gabor Kovacs and Horst Olschewski

6. The intensive care unit 85
Donald M. Arnold and Deborah J. Cook

Outcomes in clinical trials: specific outcomes

7. Biomarkers 90
Jonas Schupp and Antje Prasse

8. Health status measurement 96
Paul W. Jones

9. Electronic patient-reported outcomes 105
Sonya Ludmilla Eremenco

10. Physical activity 117
Henrik Watz

11. Lung function 127
Denis E. O'Donnell, Pierantonio Laveneziana and J. Alberto Neder

12. Imaging 146
David G. Parr

13. Exercise tests 170
Thierry Troosters

14. Omics 179
Daniel J. Kass and Naftali Kaminski

CME credit application form 189

C O P E COMMITTEE ON PUBLICATION ETHICS

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics.

Preface

The Guest Editors describe, in their introduction to this *European Respiratory Monograph* (*ERM*), a number of reasons why fewer and fewer new drugs receive regulatory approval, although the funds invested by the pharmaceutical industry for drug development are higher than ever before. Two main reasons are mentioned.

First, the phase II and III studies, which are essential for final approval by the regulatory authorities, are extremely expensive and there is no guarantee that the development costs can be recouped after marketing authorisation. Outcome parameters of early clinical trials that demonstrate a high reliability for predicting results in phase III studies are important. However, the classical outcome parameters, such as lung function measurement and symptom scores, have failed in this regard.

Secondly, drug development has stalled because targets for development are lacking. New trial outcome parameters can contribute to new target identification for therapy: targets that had not been thought of when using the classic method of drug development.

I want to congratulate the Guest Editors, Martin Kolb and Claus F. Vogelmeier, for setting up this extraordinary issue of the *ERM*. For a topic like this, most of the authors did not have a template, and the collation of the data involved substantial work and was sophisticated. Due to time constraints, it was not possible to cover the entire field of pulmonary medicine, and there are additional areas of clinical trial outcomes that might have been included in this issue if preparation time had been available.

Nevertheless, this issue of the *ERM* should be of interest for a wide readership, including basic researchers and clinicians in the field of pneumology and also people working in the pharmaceutical industry or in regulatory authorities dealing with drug approval. I am convinced that they will find this Monograph useful for further consideration of clinical trials.

Eur Respir Monogr 2013; 62: v. Copyright ERS 2013. DOI: 10.1183/1025448x.10011513. Print ISBN: 978-1-84984-044-6. Online ISBN: 978-1-84984-045-3. Print ISSN: 1025-448x. Online ISSN: 2075-6674.

Guest Editors

Martin Kolb

Martin Kolb is a German–Canadian physician who obtained his MD at the Julius-Maximilian University Medical School in Würzburg, Germany. He received extensive training in clinical and anatomical pathology in Nürnberg, Germany, before he enrolled in general internal medicine and intensive care. He specialised in respiratory medicine and was board-certified for this discipline and for internal medicine in 2002, then received his PhD equivalent the following year. Currently, Martin Kolb is Associate Professor of Medicine at McMaster University, Hamilton, ON, Canada, as well as Director of the Division of Respirology and Research Director of the Firestone Institute for Respiratory Health, also in Hamilton. He looks after several hundred patients with interstitial lung disease in his specialty clinic and also practises in general respirology. He is on the medical staff at St Joseph's Healthcare Hamilton, for respirology and general internal medicine.

In addition, Martin Kolb is Deputy Editor for *Respirology*, the official journal of the Asian Pacific Society of Respirology and Academic Editor for *PLoS One*. He is on the Editorial Boards of the *American Journal of Respiratory and Critical Care Medicine*, the *European Respiratory Journal* (*ERJ*) and the *European Respiratory Review*. Martin Kolb is on the Executive Committee for the Respiratory Cell and Molecular Biology Assembly of the American Thoracic Society (ATS) and Chair of the Lung Injury and Repair Group of the European Respiratory Society (ERS).

Martin Kolb's major research area is focused on mechanisms of lung injury, repair and fibrosis, particularly in idiopathic pulmonary fibrosis. He has a strong interest in growth factor biology, extracellular matrix and mesenchymal cell progenitors. In his lab he uses a variety of animal models to study disease mechanisms and also the efficacy of novel drugs in the preclinical setting. In addition, he leads activities in biomarker development for lung fibrosis and he participates as Principal Investigator and Steering Committee member in numerous clinical trials on interstitial lung disease. Martin Kolb has over 80 peer-reviewed publications and has received funding from the Canadian Institutes of Health Research (CIHR), the National Institutes of Health, the Canadian Foundation for Innovation, the Ontario Thoracic Society and others. He has received career awards from the Parker B. Francis Families Foundation, the Dept of Medicine at McMaster University and the New Investigator Award from the CIHR.

Eur Respir Monogr 2013; 62: vii–viii.
Copyright ERS 2013.
DOI: 10.1183/1025448x.10011313
Print ISBN: 978-1-84984-044-6
Online ISBN: 978-1-84984-045-3
Print ISSN: 1025-448x
Online ISSN: 2075-6674

Claus F. Vogelmeier

Claus F. Vogelmeier is Professor of Medicine and Head of the Dept for Pulmonary Medicine at the Philipps University of Marburg, Germany. After qualifying in medicine from the University of Munich, Claus Vogelmeier started his professional career at the hospital of the University of Munich. He is board-certified in internal medicine, pulmonary medicine, cardiology and allergology. Claus Vogelmeier spent two years as a postdoctoral fellow at the Pulmonary Branch of the National Heart, Lung and Blood Institute, Institutes of Health, in Bethesda, MD, USA. He was nominated Professor of Medicine at the Philipps University of Marburg in 2001.

Claus Vogelmeier is an active member of several respiratory societies, including the ATS and the ERS. From 2002 until 2008 he was Section Editor of the *ERJ*. From 2009 until 2011 he was President of the German Respiratory Society. In 2009 he also became Chairman of the German Asthma and Chronic Obstructive Pulmonary Disease Network. Since 2010, he has been a member of the Science and the Writing Committees of the Global Initiative for Chronic Obstructive Lung Disease (GOLD).

Claus Vogelmeier has a long-standing scientific and clinical interest in obstructive lung diseases, with topics ranging from pathogenetic aspects to novel diagnostic methods and clinical studies.

Introduction

Martin Kolb[*] *and Claus F. Vogelmeier*[#,¶]

[*]*Dept of Medicine, Pathology and Molecular Medicine, Firestone Institute for Respiratory Health, McMaster University, Hamilton, ON, Canada.* [#]*Dept of Medicine, Pulmonary and Critical Care Medicine, University Medical Centre Giessen and Marburg, Philipps-University Marburg, Marburg, Germany.* [¶]*Member of the German Centre for Lung Research.*

Correspondence: M. Kolb, Dept of Medicine, Pathology and Molecular Medicine, Firestone Institute for Respiratory Health, McMaster University, St Joseph's Healthcare, 50 Charlton Ave East, T2131, Hamilton, ON, L8N 4A6, Canada. Email: kolbm@mcmaster.ca

The last 50 years have seen a constant decline in approval rates for new therapeutic compounds with respect to the money spent in research and development across all medicine subspecialties. Almost 50% of the costs of developing new drugs are incurred during the clinical stages of the development, particularly in phase II and III trials. Therefore, many pharmaceutical companies aim to reduce expenses for these costly steps in drug development. One method of attempting this is to focus on proof-of-concept studies, which allow them to obtain either ''quick wins'' or ''fast fails'' of new products. When ''wins'' are only achievable over a prolonged period of time, rapid failures are much more cost effective. For some of the chronic lung disease areas, *e.g.* idiopathic pulmonary fibrosis and pulmonary hypertension, the newer clinical trials involve large phase II trials that have enrolled sufficiently high numbers of patients to allow a meaningful efficacy analysis. They also use the design of two parallel arms in one phase III trial to get independent study results for faster regulatory approval. In order for these new strategies in clinical research and drug development to be successful, the outcomes used for clinical trials are increasingly critical.

While the traditional end-points for clinical studies of lung diseases have been based on functional parameters, there are concerns that this method may oversimplify the complexity of diseases, particularly in the context of chronic lung disease. The value of functional outcomes as surrogate markers for disease activity and progression has been increasingly questioned by scientists, carers, regulatory agencies and funding bodies. Novel tools and methods to measure and quantify biomarkers and patient-reported outcomes have allowed these parameters to emerge from their status as interesting secondary end-points and become potential primary outcomes for clinical trials. Wireless technology and advances in electronic devices make electronic patient-reported outcomes and e-health an option, not only for improved management of patients, but also for better clinical research. Nevertheless, the relevance and validity of these novel outcomes and tools still need to be proven.

This issue of the *European Respiratory Monograph* describes the current status of commonly used as well as investigational end-points in all relevant areas of pulmonary medicine. It includes comprehensive chapters on the most prevalent respiratory disorders, from asthma to chronic obstructive pulmonary disease, lung fibrosis, cystic fibrosis and critical care problems. In the second half of this book, we provide a detailed description of specific outcomes, their value, their limitations and the future promise they might hold. We have gathered a high-profile group of authors from the international societies who are all experts in their respective fields and provide an up-to-date presentation of outcomes in clinical trials in pulmonary medicine. As editors, we hope that this issue will provide a useful summary of the state of the art to the interested reader, and form the basis for seeking additional information from the extensive literature cited in each chapter.

Eur Respir Monogr 2013; 62: ix. Copyright ERS 2013. DOI: 10.1183/1025448x.10011413. Print ISBN: 978-1-84984-044-6. Online ISBN: 978-1-84984-045-3. Print ISSN: 1025-448x. Online ISSN: 2075-6674.

Chapter 1

Chronic obstructive pulmonary disease

Emiel F.M. Wouters

SUMMARY: Definitions of chronic obstructive pulmonary disease (COPD), put forward in consensus reports and guidelines, reflect the dominance of physiological outcome parameters in clinical COPD trials. The focus on lung function changes and airflow limitation in particular seems paradoxical since, in many definitions, irreversibility of airflow limitation is part of the pathophysiological characteristics of COPD. This chapter reviews a wide variety of markers and outcomes applied in COPD intervention studies.

Dept of Respiratory Medicine, Maastricht University Medical Center, Maastricht, The Netherlands.

Correspondence: E.F.M. Wouters, Dept of Respiratory Medicine, Maastricht University Medical Center, P.O. Box 5800, 6202 AZ Maastricht, The Netherlands. Email: e.wouters@mumc.nl

Eur Respir Monogr 2013; 62: 1–18.
Copyright ERS 2013.
DOI: 10.1183/1025448x.10044212
Print ISBN: 978-1-84984-044-6
Online ISBN: 978-1-84984-045-3
Print ISSN: 1025-448x
Online ISSN: 2075-6674

At the beginning of the 21st century, chronic obstructive pulmonary disease (COPD) has taken the place occupied by tuberculosis in the first half of the 19th century. In 1990, the disease ranked twelfth in the causes of global disease burden, and it is expected to rise to fifth by 2020.

In general, the term ''disease'' is defined as a condition or state of a group of persons who have specified characteristics by which they differ from the norm in a way that is biologically disadvantageous. The name of the disease should refer succinctly to the aetiology of the disease or the abnormal phenomena displayed by the affected group of persons [1].

Two landmark meetings, the CIBA symposium in 1959 and the American Thoracic Society (ATS) Committee on Diagnostic Standards in 1962, defined the components of COPD, which are still the foundation for our definitions today [2, 3]. Chronic bronchitis is defined in clinical terms; the following condition needs to be fulfilled: chronic cough for at least 3 months during a period of at least 2 years. By contrast, emphysema was defined in anatomical terms of enlarged alveolar spaces and loss of alveolar walls. Neither definition used any physiological criteria [2, 3].

Our ability to measure lung function in the clinic and in the research laboratory, beginning in the mid-20th century, has had a profound effect on the definitions of COPD. Forced expiratory volume in 1 s (FEV_1) was established as a reliable measure of ventilatory function and the ratio FEV_1/forced vital capacity (FVC) as a measure of airflow limitation. Airflow limitation became the dominant physiological characteristic of COPD patients. In 1995, the ATS defined COPD as a disease state characterised by the presence of airflow obstruction due to chronic bronchitis or emphysema; the airflow obstruction is generally progressive, may be accompanied by airway hyperreactivity and may be partially reversible [4]. This definition uses the combination of chronic bronchitis in clinical terms, emphysema in anatomical terms and airflow obstruction representing a physiological state.

The 1995 definition from the European Respiratory Society (ERS) defined COPD as a disorder characterised by reduced maximum expiratory flow and slow forced emptying of the lungs, features which do not change markedly over several months. Most of the airflow limitation is

slowly progressive and irreversible. The airflow limitation is due to varying combinations of airway disease and emphysema but the relative contribution of these two processes is difficult to define *in vivo* [5]. Interestingly, while the initial ATS compound definition describes three independent entities, airflow limitation becomes the dominant characteristic in the ERS definition, and airway disease and emphysema are combined in a variable relative contribution to explain the airflow limitation; airflow limitation may be due to loss of elastic recoil or obstruction of the conducting airways, both small and large. In 1997, The British Thoracic Society defined COPD as a chronic, slowly progressive disorder characterised by airway obstruction (FEV_1 <80% predicted and FEV_1/FVC <70%) that does not change markedly over several months. The impairment of lung function is largely fixed but is partially reversible by bronchodilator (and other) therapy [6].

Although staging of a disease should ideally be based on a combination of factors including symptoms, severity of airflow limitation, degree of blood gas abnormalities and systemic disease manifestations, staging of severity of COPD in these different documents was largely based on the degree of airflow limitation: different cut-off values were applied for staging of the severity of COPD.

The irreversibility of change significantly determines the chronicity of the disease condition: irreversibility can take the form of continuous symptoms but also of irreparable loss of lung function. Irreversibility is in fact the product of lung function technology: it could not exist without the possibility of measuring the expiration force. Chronicity in most of these definitions is typically a gradual phenomenon and limits of reversibility are rather vague. Intriguingly, medications such as bronchodilators are not only medications to relieve symptoms, they reorganise the disease conditions by introducing the degree of reversibility as a new characteristic of the disease.

In the beginning of the 21th century, inflammatory changes were introduced as part of the definitions of COPD. In 2001, a consensus report was released by the Global Initiative for Chronic Obstructive Lung Disease (GOLD). Based on physiology, aetiology and pathology, COPD was defined as a disease state characterised by airflow limitation that is not fully reversible. The airflow limitation is usually both progressive and associated with an abnormal inflammatory response of the lungs to noxious particles or gases. In the report it is stated that the chronic inflammation produces remodelling and narrowing of the small airways as well as destruction of the lung parenchyma due to loss of alveolar attachments to the small airways and loss of elastic recoil [7]. Airflow-derived cut-off values were still put forward for staging the severity of COPD in this first GOLD document.

In 2004, a joint ERS/ATS consensus report on standards for the diagnosis and management of COPD was published. In that document, COPD was considered as a preventable and treatable disease state and it was concluded that, although COPD affects the lungs, it also produces systemic effects. In that report, it was stipulated that assessment of severity must be based on FEV_1 measurement, rating of dyspnoea and body mass index (BMI) [8].

In the following years, revisions of the original GOLD consensus report were published based on progress in our understanding of COPD. In the latest 2013 revision, GOLD defined COPD as a common preventable and treatable disease, characterised by persistent airflow limitation that is usually progressive and associated with an enhanced inflammatory response in the airways and the lung to noxious particles or gases. Exacerbations and comorbidities contribute to the overall severity in individual patients [9]. In the latest revision, GOLD advised a combined COPD assessment based on symptoms, risk of exacerbations and the degree of airflow limitation to assess the severity of COPD. Furthermore, GOLD formulates that comorbidities can occur in every COPD patient, independently of the degree of airflow limitation, and that these comorbidities influence mortality and hospitalisations independently [9–11].

This description of COPD definitions put forward in consensus reports during the last two decades is important for the understanding of the dominance of physiological outcome parameters in many COPD clinical trials. In the absence of an aetiology-based definition, such as for diseases caused by infectious agents, altered structure or function had the highest priority and

clinical features of the disease were largely neglected. The current recognition that COPD is a syndrome with a wide spectrum of respiratory and multi-organ manifestations offers new opportunities to identify underlying aetiological processes and to define particular disease entities now shared under the COPD umbrella. α_1-Antitrypsin deficiency can be considered as such an aetiology-based disease.

In the absence of aetiological or mechanistic interventions, consensus reports put forward symptomatic treatment as the most important management goal in stable COPD patients. In the latest GOLD revision, reduction in symptoms is combined with a reduction of future risks for the COPD patient. Reduction of symptoms is operationalised in reduction in breathlessness, improvement in exercise tolerance and improvement in health status. Prevention of disease progression and exacerbations as well as reduction in mortality are treatment goals in order to achieve a reduction of risks. These goals are assumed to be largely achieved by pharmacological treatment, consisting predominantly of bronchodilating β_2-agonists and anticholinergic agents and anti-inflammatory corticosteroid treatment.

Outcomes and markers

COPD is now considered to be a heterogeneous clinical syndrome reflecting the interaction of a complex range of pathological changes including both pulmonary and systemic effects. In the context of measurement of the effect of a disease or the effect of interventions, outcomes and markers are used. In a review by JONES and AGUSTI [12], "outcome" is restricted to the consequences of the disease as experienced by the patient. In COPD, a number of patient-related outcomes are recognised as being important, such as symptoms, dyspnoea, functional status and impaired exercise tolerance, health-related quality of life and health status, exacerbations and death. These patient-related outcomes are global or summative outcomes: they have the potential to aggregate multiple small effects or the overall effect of a number of biological steps [13, 14]. These global outcomes can demonstrate that a change has occurred and provide an assessment of the clinical significance of the effect. The mechanisms behind the observed summative and aggregated effects remain unidentified. Therefore, these global or summative outcomes are considered very interesting for hypothesis generation [13].

A marker is a measurement known to be associated with a clinical outcome [12]. In the absence of widely accepted and validated markers of COPD, FEV1 has been used as a global marker of all the pathophysiological changes in COPD, although measures of lung function correlate poorly with patient-related outcomes [15]. Considering the diverse range of pathological mechanisms involved in COPD, it is very unlikely that a single marker such as FEV1 provides a full assessment of the impact of the disease. Based on the historical developments of COPD study, the role of FEV1 as a global marker of COPD is reinforced by the requirements of regulatory authorities that treatments should improve FEV1. As a consequence, treatments with the potential to modify patient-related outcomes by nonbronchodilating actions might be denied to physicians and patients, based on these regulatory requirements [16]. The focus on FEV1 as the single global marker to assess treatment efficacy seems paradoxical as, even by definition, airflow limitation is irreversible in COPD. Considering that all COPD definitions include progression of airflow limitation, the rate of decline in FEV1 over time is considered as a marker of disease progression and a target for pharmacological interventions.

Decline in FEV1 used as a marker of disease progression in clinical trials

The natural history of tobacco smoke-related chronic airflow obstruction is largely based on the landmark study of FLETCHER and PETO [17]. In their pioneering work, they demonstrated that the rate of FEV1 decline increased progressively with time in susceptible smokers. Further conclusions

derived from this landmark study and still dominating current COPD practice are the findings of individual susceptibility to smoking-related lung function damage, unavoidable progression of airflow obstruction once COPD is established, large variability of abnormal lung function decline among different subjects, possible reduction of the excessive FEV1 decline after smoking cessation at all times, and the paucity of self-reported symptoms when airflow limitation is mild to moderate [17, 18]. Modified Fletcher and Peto graphs, summarising these data, are still used in COPD educational programmes. The many limitations of the study are largely overlooked in the literature. One of these is that the study included only 792 working males, followed every 6 months for an 8-year follow-up period. Despite this short follow-up period, the x-axis on the graphs of FLETCHER and PETO [17] describes the changes in FEV1, expressed as the percentage of the value at age 25 years, between 25 and 75 years. Therefore, the graphs are largely the result of extrapolations backward and forward with inherent inaccuracies as the recruitment age was between 30 and 59 years only [17].

Different studies have revisited the natural history of chronic airflow obstruction as the dominant physiological characteristic in COPD. All these studies confirm that tobacco smoking increases the rate of lung function decline in males and females [19–27]. In contrast to the data of FLETCHER and PETO [17], more recent studies indicate that the presence of respiratory symptoms at baseline and/or a respiratory diagnosis during follow-up appears to identify a subpopulation of smokers who are susceptible to the deleterious effects in their lungs of smoking [27, 28]. More recent studies also highlight the importance of early quitting and do not support the model of FLETCHER and PETO [17], hypothesising that the rate of lung function decline returns to normal after quitting, independent of the age of the individual [29, 30].

Based on an analysis of the placebo arms of large observational studies with a follow-up of at least 3 years, a faster progression of functional impairment is reported in the earlier stages of COPD than in the later stages. Based on these data, therapeutic options should be focused on COPD patients in at least GOLD stage II [18].

As part of the ECLIPSE study (Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints), the changes in post-bronchodilator FEV1 were analysed over a period of 3 years to assess the variability as well as the determinants of changes in FEV1. It was found that the rate of change in FEV1 was highly variable: a quarter of the patients had a change that ranged from a decrease of 20 to an increase of 20 mL·year^{-1}, and 8% had an increase of more than 20 mL·year^{-1}. Increased rates of decline were reported in current smokers, patients with bronchodilator reversibility and patients with emphysema [31]. At the least, these data show that COPD is not invariably progressive.

Many prospective studies have tested the effects of different pharmacological interventions on the slope of FEV1. Inhaled short-acting anticholinergic drugs, inhaled corticosteroids and *N*-acetylcysteine did not alter the decline of FEV1, used as a marker of disease progression [32–38]. A study with tiotropium bromide was specifically designed to evaluate the effects of this drug on the rate of decline in FEV1 over a period of 4 years. Therapy with tiotropium did not significantly reduce the rate of decline in FEV1 [39]. In a pre-specified subgroup analysis in COPD patients with moderate airflow limitation (GOLD II), an attenuated decline of mean post-bronchodilator FEV1 was reported [40]. In a *post hoc* analysis of the TORCH study (Towards a Revolution in COPD Health), the three active treatment arms (salmeterol, fluticasone and the combination of salmeterol and fluticasone) reduced the rate of FEV1 decline (13–16 mL·year^{-1}), but rate of decline was similar among these three active treatment arms [41]. It must be realised that in the absence of an authentic intent-to-treat analysis, results can be biased from regression to the mean [42].

In prospective studies, only smoking cessation has been shown to alter the rate of decline in FEV1 in patients with COPD [33]. Future studies are needed to identify other markers reflecting progression of pathophysiological processes involved in the multicomponent pathology of COPD.

FEV1 as a specific physiological marker

Specific markers are assumed to measure a single biological variable and address a unidimensional construct such as airway obstruction. FEV1 is the volume of air that is forcibly exhaled in the first second, whereas FVC is the total volume of air exhaled after a full inspiration. FEV1 is a summative measure in COPD because it reflects both disease in the airway wall as well as the loss of alveolar attachments caused by emphysema [13].

The methodology for obtaining forced expiratory manoeuvres and derived parameters has been standardised by the ATS and ERS [43]. Reference equations based on the distribution of spirometric parameters are available for different ethnic groups and age ranges [44–46]. An important limitation is the lack of reference values in the normal elderly population. FEV1 and FVC parameters have been shown to be highly reproducible if obtained by well-trained technicians. The ATS/ERS recommendations consider a change in FEV1 >20% in short-term trials and >15% in long-term trials as clinically meaningful [46]. The European Medicines Agency specifies strict spirometric and reversibility criteria to patient inclusion in clinical trials [16].

FEV1 has been used as a primary variable in the vast majority of multicentre trials related to COPD, even for interventions with pharmacological agents without pharmacological bronchodilating properties [47–49]. The use of surrogate end-points such as FEV1 is most understandable for interventions aimed to induce airway smooth muscle relaxation. These bronchodilating drugs (β-agonists, anticholinergic derivates) form the cornerstone of current symptomatic management of COPD. Besides possible changes in mucociliary clearance by these drugs and assumed anti-inflammatory actions, it is important to realise that the effects on measured FEV1 are a summation of changes in the central and/or peripheral airways, resulting in reduction of the inspiratory and expiratory airways resistance and a reduction in the respiratory system time constants. Furthermore, the complex mechanical consequences of bronchodilatation are frequently ignored. Bronchodilators promote a slower, deeper breathing pattern, favouring more complete lung emptying. These changes will reduce the work of breathing due to hyperinflation and reduce the threshold load of intrinsic positive end-expiratory pressure, resulting in an increased inspiratory force reserve [50].

Based on a better understanding of the respiratory mechanics, it is clear that relatively modest changes in airway dimensions can have major respiratory mechanical effects as well as cause important improvements in respiratory symptoms. These effects can be missed by application of a single physiological marker and can lead to an unduly pessimistic view of the benefits of certain drugs. Even for bronchodilating drugs, it is recommended that a number of physiological markers are used, instead of the current simplified evaluation.

Markers of anti-inflammatory effects

As an enhanced inflammatory response to noxious particles or gases is currently part of the definition of COPD, it is not surprising that suppression or elimination of inflammation has been extensively studied as part of COPD management. Most attention has focused on suppression of inflammation by corticosteroids, administered orally or by inhalation, as these drugs have marked anti-inflammatory effects in asthma. However, in contrast with asthma, inflammation is still poorly defined in COPD. It remains unclear whether quantitative or qualitative differences in the inflammatory response account for the development of COPD in susceptible subjects. Historically, the neutrophil was considered as a logical candidate and the potential to release elastase and proteolytic enzymes was linked to the development of alveolar destruction. Current studies conclude that the presence of pulmonary neutrophilia does not seem to differentiate between COPD patients and smokers [51–54]. The finding of increased numbers of T-lymphocytes and especially CD8+ T-cells in smokers who develop COPD favours the notion that tissue injury is dependent on T-cell activation [55, 56]. In COPD patients, airway inflammation seems very heterogeneous and poorly linked to the degree of airflow limitation [57].

Limited and inconclusive data are reported about the anti-inflammatory effects of inhaled steroids. In one study, a significant reduction in the number of mucosal mast cells was accompanied by an increase in neutrophils in the bronchial biopsies [58]. Another study could not find any reduction in the primary end-points of CD8+ cells, CD68+ cells or neutrophils [59]. The authors reported a reduction in the CD8:CD4 ratio in the epithelium, largely related to the increase in CD4+ cells in both the subepithelium and epithelium. In the same study, an increase in the number of subepithelial mast cells was reported in the group of patients treated with inhaled steroids [59]. Another study reported a reduction in macrophage number but an increase in neutrophil number in endobronchial biopsies after inhaled steroid therapy, as well as reductions in the number of mast cells and CD8+ lymphocytes after inhaled steroids [60]. In contrast, BOURBEAU *et al.* [61] could not find any marked effect on CD8+ cells or CD68+ macrophages after therapy with inhaled steroids alone.

These studies have assessed anti-inflammatory effects after a short period of intervention (3–6 months). Recently, an intervention study was conducted with bronchial biopsies at 6 and 30 months [62]. In that study, the authors reported decreased counts of mucosal CD3+ cells, CD4+ cells, CD8+ cells and mast cells after 6 months of intervention with inhaled steroids. Fluticasone therapy for 30 months reduced mast cell count and increased eosinophil count and percentage of intact epithelium, with accompanying reductions in sputum neutrophil, macrophage and lymphocyte counts. Discontinuation of inhaled steroids increased the counts of CD3+ cells, mast cells and plasma cells [62]. Combination therapy with a long-acting β_2-agonist, salmeterol, was associated with a reduction in biopsy CD8+ cells with unaffected counts of CD68+ cells [63]. The combination therapy also significantly reduced biopsy CD45+ and CD4+ cells expressing genes for tumour necrosis factor (TNF)-α and interferon-γ [62]. Others reported reductions of CD8+ cells and CD68+ macrophages after combination therapy with salmeterol and fluticasone only [61].

The limited number of studies and studied patients, the short duration of interventions and the conflicting outcomes contribute to the existing controversy about the so-called anti-inflammatory effects of inhaled steroids in COPD. Even a systematic review including 55 studies could not formulate a firm conclusion [64]. In that study, rate of decline in FEV_1, mortality, mean rate of exacerbations and rate of decline in health status were assessed as summative or global outcomes to assess effects of inhaled steroids: reduced rate of exacerbations and reduced rate of the decline of health status were identified as significant benefits of this pharmacological intervention [64]. Intriguingly, the efficacy of specific anti-inflammatory interventions, such as anti-TNF or interleukin-8 antagonists, is studied in COPD patients using summative outcomes such as health status, dyspnoea sensation or the number of COPD exacerbations [65]. Smaller pilot studies indicate that these biological interventions do not modify local and/or systemic inflammation [66, 67].

Imaging

Computed tomography (CT) imaging provides a means of accurately characterising lung parenchymal changes and the nature of the image data facilitates quantitative assessment. CT is more sensitive than plain radiography in diagnosing emphysema and correlates with the presence and severity of emphysema better than nonspecific physiological tests such as FEV_1 and single-breath carbon monoxide diffusing capacity [68, 69]. Various indices have been used to quantify lung density changes by CT and the airway wall thickening can be evaluated in the conducting airways.

Serial CT scanning has been proposed as a potential method to determine the progression of emphysema and trials of α_1-antitrypsin augmentation therapy have used change in CT lung density as an outcome parameter [70–72]. Limited data are available about the change in lung density in COPD patients. As part of the ECLIPSE study, lung density was followed-up in 1928 COPD patients over 3 years. Lung density was expressed as the volume-adjusted 15th percentile density. It was found that the observed decline in lung density was variable and resulted in a mean

decline of 1.13 $g \cdot L^{-1} \cdot year^{-1}$. This progression in emphysema varied substantially between subjects: females had a more rapid decline than males, as did current smokers *versus* former smokers, and patients with a higher baseline BMI were associated with a reduced decline in lung density. Based on the results of that study, the current value of CT in clinical trials seems to be more of a stratification tool [73]. Lung densitometry may become a future marker of COPD, particularly if more specific therapeutic interventions are available to tackle these structural abnormalities.

Exacerbations as an outcome

In many clinical trials, exacerbations are integrated as an outcome measure based on the negative impact of exacerbations for COPD patients [74–77]. Furthermore, acute exacerbations are the most frequent cause of medical visits, hospital admissions and death among patients with COPD [78]. One of the main objectives in clinical trials is to reduce the frequency and severity of these exacerbations. These studies are hampered by the absence of a unanimous definition of exacerbations of COPD, the absence of criteria to define the severity of exacerbations and the neglect to recognise that exacerbations are very heterogeneous in clinical presentation and underlying biological processes and no set of validated markers or biomarkers for acute COPD exacerbations is available.

At present, four definitions of exacerbation are still widely used. The first definition uses a combination of three cardinal symptoms: increased dyspnoea, sputum volume and sputum purulence. Depending on the number of reported cardinal symptoms, exacerbations are further subdivided into three categories (I-III) [79]. This definition of acute exacerbations has been widely used in clinical trials of antibiotics for acute exacerbations of COPD. The second definition of exacerbations is based on the presence of the following patterns of symptoms during two or more consecutive days: either two or more of the three major symptoms (increase in dyspnoea, sputum volume or sputum purulence); or any one major symptom together with any one minor symptom (increase in nasal discharge, wheeze, sore throat, cough or fever) [74]. This definition has been used in follow-up studies of COPD patients and offers the advantage that all exacerbations defined according these criteria can be identified, independently of whether they are reported to the healthcare professionals. The third definition is a consensus definition of an expert panel: a sustained worsening of the patient's condition from the stable state and beyond normal day-to-day variations, which is acute in onset and necessitates a change in regular medication in a patient with underlying COPD [80]. The fourth definition identifies exacerbations as a complex series of respiratory events (cough, wheezing and dyspnoea) lasting three or more days [37, 81].

All these definitions are still poorly validated and largely based on symptoms reported by the patient. Some of these symptoms are related to underlying aetiological conditions, such as viral or bacterial infections. Remarkably, while COPD itself is largely based on physiological criteria, exacerbations, the major driver of disease burden, are only defined in qualitative, descriptive terms.

Most interventions focus on reduction of the frequency and severity of these exacerbations. However, no validated scale of severity exists for exacerbations. Most studies have used the intensity of the medical intervention required as a grade of severity [82, 83]. The proposed severity classification includes three categories: 1) mild, which involves an increase in respiratory symptoms that can be controlled by the patient with an increase in usual medication; 2) moderate, which requires treatment with systemic steroids and/or antibiotics; and 3) severe, which describes exacerbations that require hospitalisation or a visit to the emergency department. Importantly, the methodology of this severity classification is not standardised at all. Furthermore, the use of systemic steroids and/or antibiotics in the management of exacerbations is still questionable [64] and has different patterns of prescription in different areas. In addition, the criteria for hospital admission may vary from country to country and even between hospitals. Criteria for hospital admissions can be unrelated to the disease condition itself, as are age of the patient or social factors necessitating hospital admission. From this perspective, the current severity classification of exacerbations can be considered as a summative outcome combining disease- and patient-related factors.

The frequency of exacerbations is also considered an important outcome for the clinical history of COPD. However, we know little about the incidence and determinants of acute exacerbations at various levels of disease severity. The frequency of exacerbations is largely determined by the definition and recording methods used. In observational studies of COPD patients, a skewed distribution of the variable has been found, with a large number of patients having 0–2 exacerbations per year [75, 84, 85]. The mean number of exacerbations in these studies is generally related to the severity of the baseline disease [75, 84, 85]. This observation was confirmed in a 3-year follow-up study: exacerbation rates in the first year of follow-up were 0.85, 1.34 and 2.00 per person for patients with GOLD stage II, III and IV, respectively [86]. Overall, 22%, 33% and 47% of patients with stage II, III and IV, respectively, had frequent exacerbations [86]. In this study, moderate exacerbations were classified based on the decision to prescribe antibiotics or systemic corticosteroids, alone or in combination. Moderate or severe exacerbations, requiring hospitalisation, were combined. Exacerbations were significantly associated with worsening lung function, greater impairment in health status, a history of gastro-oesophageal reflux and an increased white blood cell count. The so-called frequent-exacerbation and no-exacerbation phenotypes appeared to be relatively stable over a period of 3 years. The single best predictor of exacerbations in this study was a history of exacerbations according to healthcare utilisation criteria [86]. Based on these data, the concept of an independent susceptibility phenotype of COPD patients is introduced and integrated in the current guidelines [9]. Again, healthcare utilisation criteria are translated to disease-related characteristics and underlying biological processes are assumed.

The lack of a clear definition describing the acute-on-chronic pathology occurring in the natural history of COPD, which is defined as a multicomponent disease condition, as well as the lack of validated scoring systems to assess severity, hampers the development of exacerbation-related intervention studies. Studies must focus on identification of markers related and associated with acute exacerbations of COPD.

Exercise

In chronic cardiopulmonary diseases the ability to exercise can be considered as an important clinical outcome and as a marker of other significant outcomes. In COPD patients, exercise tolerance cannot be predicted by resting lung function measurements and exercise testing is useful to assess the degree of impairment, prognosis and effects of interventions. Exercise tolerance can be limited by ventilatory constraints, lung gas exchange abnormalities, lung dynamic hyperinflation and subjective symptoms such as dyspnoea sensation, as well as by deconditioning and leg fatigue [87]. The severity and cause of exercise tolerance can be assessed by performing detailed physiological measurements in the laboratory and so-called field tests, where the duration of exercise or the distance covered in a fixed period can be measured.

Laboratory exercise testing is conducted under standardised conditions, with recording of cardiac frequency, minute ventilation, respiratory frequency and breathing pattern, oxygen uptake and carbon dioxide production, exercise duration and maximum workload performed. Lung volumes are generally measured by measurement of the inspiratory capacity. The intensity of respiratory sensation, breathlessness and leg fatigue can be recorded.

The introduction of the concept of dynamic hyperinflation as one of the major limitations to exercise performance in COPD, and data reporting the mismatch in COPD patients between the central neural drive and the abnormal restricted mechanical response as one of the fundaments in the origin of dyspnoea, have contributed to a need for intervention studies aiming to modify lung hyperinflation [88, 89]. Initially reported to be one of the outcomes of long-acting anticholinergic interventions [90, 91], it is now recognised that all classes of bronchodilators that act by relaxing airway smooth muscle tone decrease the mechanical time constants for emptying in heterogeneously distributed alveolar units [92]. Sustained increases in the resting inspiratory capacity of around 0.3 L appear to be clinically meaningful [91, 93–96]. In selected groups of patients, lung volume reduction surgery and bullectomy have been shown to have similar

improvements in operating lung volumes, effort-displacement ratios, exertional dyspnoea and exercise performance [97, 98]. Modifying the physical properties of inhaled air and exercise training also modify the rate of dynamic hyperinflation during exercise [96, 99, 100].

Field tests can be distinguished in self-paced timed walking tests and shuttle walking tests. In self-paced walking tests, patients are asked to walk as far as possible for a set time, originally 12 min but now usually 6 min [101, 102]. Patients are allowed to stop if they cannot continue, but are asked to resume walking when they are able to. The principal outcome is the distance walked. The intensity of respiratory sensation, oxygen saturation or cardiac frequency can be recorded during these walking tests. These tests are reproducible in both respiratory and cardiac patients [103, 104].

The 6-min walking distance (6MWD) is an important predictor of mortality and health status in COPD patients [105–107]. Given the prognostic significance of the 6MWD, it is important to identify determinants of this functional performance parameter. Many factors can influence 6MWD in COPD, including sex, body composition, degree of airflow limitation, degree of emphysema and the rate of acute exacerbations [108–112]. Furthermore, perceptions of poor health correlate with the 6MWD in COPD [113–115]. A recent study confirmed that both physical and psychological factors are important determinants of COPD: more severe airflow limitation, the degree of emphysema, oxygen use during/after the 6-min walk test (6MWT), presence of depressive symptoms and moderate-to-severe symptoms [116].

The 6MWD is accepted as a valid tool and primary outcome in pulmonary conditions such as idiopathic pulmonary arterial hypertension and idiopathic pulmonary fibrosis [117, 118]. Studies have confirmed that multiple interventions, such as pulmonary rehabilitation, lung volume reduction surgery, oxygen during exercise and bronchodilator drugs, improve this outcome significantly [119–123].

In designing studies in which 6MWD is an outcome measure, knowledge of the minimum clinically important difference (MCID) is essential in interpreting the results. A recent study established the MCID of patients with COPD with a wide range of degrees of airflow limitation, using mortality as an anchor measure. A reduction in the 6MWD of 30 m represents a significant MCID in this population [124]. This change seems a reasonable threshold that falls between the 25 and 54 m that have been reported before [125].

As an alternative to the unpaced 6MWT, shuttle tests have been developed in an attempt to improve standardisation and reproducibility. Walking speed is externally paced by the frequency of bleeps on a pre-recorded tape. The principal outcome is the distance covered in metres [126]. The shuttle walk test has been applied in various studies and this measure seems responsive after pulmonary rehabilitation, nutritional support and bronchodilator drugs [127–130].

Besides exercise testing, interest is growing in directly or indirectly assessing physical activity itself. Physical activity can be defined as any bodily movement produced by skeletal muscles, which results in energy expenditure. Physical activity is a complex behaviour that can be characterised by intensity, duration and patterns; exercise is a subset of physical activity. Physical activity also includes, but is not limited to, leisure-time, domestic and occupational activities. In recent years, attention has been raised regarding the adverse effects of a sedentary lifestyle, characterised by behaviours that do not increase skeletal muscle energy expenditure substantially above the resting level [131, 132]. Physical activity can be assessed by questionnaires, step counters and activity monitors, as well as by stable isotopes (doubly labelled water). The use of activity monitors in particular offers interesting opportunities for integration in intervention studies, but a variety of methodological aspects need to be solved and the lack of uniformity in output from the various accelerometers makes comparison of studies difficult.

Patient-reported outcomes

A patient-reported outcome is a report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else.

Dyspnoea is a common, distressing symptom in COPD patients, but is also a common symptom in cardiopulmonary and neuromuscular diseases in general. Mechanisms, assessment and management of dyspnoea have recently been reviewed in an updated ATS statement [133]. Dyspnoea is defined as a subjective experience of breathing discomfort that consists of qualitatively distinct sensations that vary in intensity. This experience of dyspnoea derives from interactions among multiple physiological, psychological, social and environmental factors and may induce secondary physiological and behavioural responses [134]. Dyspnoea can only be perceived by the person experiencing it. Dyspnoea as a complex symptom must be distinguished from signs of respiratory distress, such as tachypnoea, use of accessory muscles and intercostal retractions.

The ATS statement proposes that instruments or sections of instruments pertaining to dyspnoea should be classified as pertaining to domains of sensory-perceptual experience, affective distress or symptom/disease impact or burden [133]. Sensory-perceptual measures evaluate what breathing "feels like" to the patient and include ratings of intensity or sensory quality. Affective distress measures how distressing breathing feels, either immediate or evaluative. Impact measures how breathing affects behaviours, beliefs or values that matter to patients or society. These measures generally involve multi-item scales, such as functional performance or disability, quality of life or health status and psycho-social functioning [133].

The Medical Research Council (MRC) scale is a unidimensional rating of disability or activity limitation. The MRC scale is a five-point scale that considers certain activities, such as walking or climbing stairs, which provoke breathlessness [135]. The MRC scale is considered a discriminative instrument for categorisation of patients in terms of disability and it has been introduced as one of the classification parameters in a recent GOLD document [9, 136]. The MRC scale is considered to be an instrument that is unresponsive to interventions and is also not a useful measure during exacerbations of COPD [137, 138].

The baseline dyspnoea index (BDI) and the transition dyspnoea index (TDI) are instruments developed to understand the patient's dyspnoea severity based on the following components: functional impairment, magnitude of task and magnitude of effort [139]. BDI quantifies the severity of dyspnoea at an initial or baseline state, whereas TDI is considered as an evaluative instrument used to quantify the changes in dyspnoea from the initial or baseline state [139]. This instrument has been applied in numerous clinical trials and in assessing the burden of exacerbations of COPD [81, 140–150]. Far more instruments for measuring dyspnoea are available. Therefore, the ATS statement formulates that this profusion of measures makes it very difficult to compare results across studies and draw evidence-based conclusions. In addition, this document stresses the need for one or more dyspnoea measures to be adequately validated as patient-reported outcomes for use as end-points in clinical trials [133].

In order to measure the disease burden and the impact of COPD on daily life, the evaluation of quality of life (QoL) has become an important outcome measure in COPD research and treatment in the last decade. QoL is usually defined as an individual's perception of their position in life or life satisfaction [151]. This perception is affected in a complex way by physical health, psychological state, level of independence, social relationships and personal beliefs [152–155]. Health status, functional status and QoL are often used interchangeably, but these constructs are not equal [151, 156, 157].

At present, many QoL instruments are available: generic QoL instruments, which can be used in any population, and so-called disease-specific instruments. These health status measurements are developed to quantify, in a standardised and objective manner, the impact of the disease on the patient's daily life, health and well-being [106]. Recognising the multiple consequences of COPD, there was growing need for an instrument that could aggregate into a single score the summed effect of the multiple primary and secondary pathophysiological processes and effects of the disease. To select an appropriate instrument for use in COPD practice or in COPD research, the measurement properties in terms of validity, reliability and responsiveness must be adequate,

instruments for evaluation of an intervention must be responsive over time and instruments for distinguishing patients must be reliable [155, 158–161].

A systematic review of the content and psychometric properties of available disease-specific and generic QoL instruments has recently been published [155]. The content of the 13 disease-specific and 10 generic instruments showed great variation: 20 instruments measured mobility; pulmonary symptoms such as dyspnoea, cough, phlegm, chest tightness and wheezing were measured in 11 disease-specific instruments. Pain, vitality and spiritual activities were domains only present in generic instruments [155]. Social and emotional functioning were domains seen in disease-specific as well as in generic instruments. The methodological quality of disease-specific questionnaires, such as the Chronic Respiratory Questionnaire (CRQ), COPD Assessment Test (CAT), St George's Respiratory Questionnaire (SGRQ) and Living with COPD Questionnaire (LCOPD), were rated positively [106, 155–157, 162]. GOLD recommends the short Clinical COPD Questionnaire (CCQ) and CAT in assessment of the severity of COPD. The CCQ consists of 10 items with an overall score, and three domains: symptoms (four items), functional state (four items) and mental state (two items). All scores range from 0 to 6 [163]. The CCQ is reported as a valid and reliable instrument for assessment of health status on the group level in patients treated for COPD in primary care, but reliability is questioned for the monitoring of individual patients [164]. The CAT is a short, 8-item, health status questionnaire developed to provide a simple tool for quantifying the overall symptomatic impact of COPD [165]. It has been shown to distinguish between different severities of COPD and is highly correlated with the SGRQ [166].

These health status instruments have found their widest application in clinical trials by providing a measure of the overall symptomatic benefit from the treatment, together with an index of whether the effect is worthwhile [106]. Health status instruments can offer a tool to evaluate the symptomatic benefits of different intervention strategies, such as pharmacological and nonpharmacological interventions (*e.g.* pulmonary rehabilitation).

QoL data can be used to help evaluate the cost-effectiveness or cost-utility of healthcare interventions. Cost-effectiveness analysis typically quantifies the benefits of a healthcare intervention in terms of years of life or quality-adjusted life-years. Combination of effectiveness and costs has increased dramatically in recent years, in order to support decision making for investments in new therapeutic interventions.

Mortality

Mortality has been the gold standard in the evaluation of therapies and interventions in different disease areas. It is reliable and of great importance in the management of patients. Improvements in mortality with COPD therapies are scantily documented: there have been two trials evaluating oxygen therapy for hypoxaemic patients and a small study evaluating the outcome of lung volume reduction surgery [167–169]. Death was also the primary outcome of a randomised trial comparing the combination of salmeterol and fluticasone with salmeterol alone, fluticasone alone and placebo [170]. The reduction in death from all causes among patients with COPD in the combination therapy group did not reach the predetermined level of significance. This study is followed-up by the SUMMIT study (Study to Understand Mortality and Morbidity in COPD), which is designed to prospectively compare the efficacy of fluticasone/vilantrol in combination, fluticasone alone and vilanterol alone, against matched placebo, on survival in subjects with moderate COPD and a history of, or an increased risk for, cardiovascular diseases [171].

All-cause mortality is probably the best COPD outcome measure and patients with advanced COPD die from a variety of other causes. Therefore, there is emerging consensus that clinical trials, evaluating death as an outcome, should have an independent adjudicating committee [172].

Conclusion

Although COPD is now recognised as a syndrome with a wide spectrum of respiratory and multi-organ manifestations, physiological markers of airflow limitation in particular remain a primary

end-point in COPD pharmacological trials. The dominance of the physiological approach in COPD management is also reflected by the many interventions aiming to modify the decline in FEV1 as an assumed marker of disease progression.

Although an enhanced inflammatory response in the airways and the lung to noxious particles or gases is part of the definition of COPD, the role of this marker in relation to clinical outcome is still debatable. Besides these markers, a wide variety of outcomes are studied in COPD intervention trials: outcomes refer to the consequences of the disease as experienced and reported by the patient. These outcome measurements have found wide application in clinical trials by providing a measure of the overall, summative symptomatic benefit of an intervention, in the absence of interventions targeting underlying mechanisms. The lack of a clear definition of outcomes such as exacerbations of COPD, which largely rely on symptoms and unvalidated scoring of severity, hampers interpretation of these outcomes. Although many pulmonary biomarkers have been described in COPD patients, the use of biomarkers for assessing new therapies in COPD still needs further research. Identification of intermediate and novel phenotypes within the clinical syndrome of COPD may provide more targeted mechanism-based interventions and will require specific markers in future pharmacological and biological interventions.

Statement of Interest

E.F.M. Wouters declares board membership of Nycomed, grants from AstraZeneca and GSK, and payment for lectures from AstraZeneca, GSK and Novartis.

References

1. Definition of asthma. *In*: Weiss EB, Stein M, eds. Bronchial Asthma, Mechanisms and Therapeutics. 3rd Edn. Boston, Little Brown and Co., 1993; pp. 1–18.
2. Terminology, definitions and classifications of chronic pulmonary emphysema and related conditions: a report of the conclusions of a CIBA guest symposium. *Thorax* 1959; 14: 286–299.
3. American Thoracic Society. Chronic bronchitis, asthma and pulmonary emphysema: a statement by the Committee on Diagnostic Standards for Nontuberculous Respiratory Diseases. *Am Rev Respir Dis* 1962; 85: 762–768.
4. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease. American Thoracic Society. *Am J Respir Crit Care Med* 1995; 152: S77–S121.
5. Siafakas NM, Vermeire P, Pride NB, *et al.* Optimal assessment and management of chronic obstructive pulmonary disease (COPD). The European Respiratory Society Task Force. *Eur Respir J* 1995; 8: 1398–1420.
6. BTS guidelines for the management of chronic obstructive pulmonary disease. The COPD Guidelines Group of the Standards of Care Committee of the BTS. *Thorax* 1997; 52: Suppl. 5, S1–S28.
7. Pauwels RA, Buist AS, Calverley PM, *et al.* Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop summary. *Am J Respir Crit Care Med* 2001; 163: 1256–1276.
8. Celli BR, MacNee W, ATS/ERS Task Force. Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. *Eur Respir J* 2004; 23: 932–946.
9. Vestbo J, Hurd SS, Agusti AG, *et al.* Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* 2013; 187: 347–365.
10. Maltais F, Ostinelli J, Bourbeau J, *et al.* Comparison of nebulized budesonide and oral prednisolone with placebo in the treatment of acute exacerbations of chronic obstructive pulmonary disease: a randomized controlled trial. *Am J Respir Crit Care Med* 2002; 165: 698–703.
11. Woodhead M, Blasi F, Ewig S, *et al.* Guidelines for the management of adult lower respiratory tract infections. *Eur Respir J* 2005; 26: 1138–1180.
12. Jones PW, Agusti AG. Outcomes and markers in the assessment of chronic obstructive pulmonary disease. *Eur Respir J* 2006; 27: 822–832.
13. Jones PW, Kaplan RM. Methodological issues in evaluating measures of health as outcomes for COPD. *Eur Respir J* 2003; 21: Suppl. 41, 13s–18s.
14. Kaplan RM. Using quality of life information to set priorities in health policy. *Soc Indic Res* 1994; 33: 121–163.
15. Curtis JR, Deyo RA, Hudson LD. Pulmonary rehabilitation in chronic respiratory insufficiency. 7. Health-related quality of life among patients with chronic obstructive pulmonary disease. *Thorax* 1994; 49: 162–170.
16. European Medicines Agency. Guideline on clinical investigation of medicinal products in the treatment of chronic obstructive pulmonary disease (COPD). London, EMA, 2012. www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/08/WC500130880.pdf

17. Fletcher C, Peto R. The natural history of chronic airflow obstruction. *Br Med J* 1977; 1: 1645–1648.
18. Tantucci C, Modina D. Lung function decline in COPD. *Int J Chron Obstruct Pulmon Dis* 2012; 7: 95–99.
19. Xu X, Weiss ST, Rijcken B, *et al.* Smoking, changes in smoking habits, and rate of decline in FEV1: new insight into gender differences. *Eur Respir J* 1994; 7: 1056–1061.
20. Tager IB, Segal MR, Munoz A, *et al.* The effect of maternal cigarette smoking on the pulmonary function of children and adolescents. Analyses of data from two populations. *Am Rev Respir Dis* 1987; 136: 1366–1370.
21. Sorlie P, Lakatos E, Kannel WB, *et al.* Influence of cigarette smoking on lung function at baseline and at follow-up in 14 years: the Framingham Study. *J Chronic Dis* 1987; 40: 849–856.
22. Mannino DM, Reichert MM, Davis KJ. Lung function decline and outcomes in an adult population. *Am J Respir Crit Care Med* 2006; 173: 985–990.
23. Robbins DR, Enright PL, Sherrill DL. Lung function development in young adults: is there a plateau phase? *Eur Respir J* 1995; 8: 768–772.
24. Prescott E, Bjerg AM, Andersen PK, *et al.* Gender difference in smoking effects on lung function and risk of hospitalization for COPD: results from a Danish longitudinal population study. *Eur Respir J* 1997; 10: 822–827.
25. Gold DR, Wang X, Wypij D, *et al.* Effects of cigarette smoking on lung function in adolescent boys and girls. *N Engl J Med* 1996; 335: 931–937.
26. Kerstjens HA, Rijcken B, Schouten JP, *et al.* Decline of FEV1 by age and smoking status: facts, figures, and fallacies. *Thorax* 1997; 52: 820–827.
27. Kohansal R, Martinez-Camblor P, Agusti A, *et al.* The natural history of chronic airflow obstruction revisited: an analysis of the Framingham offspring cohort. *Am J Respir Crit Care Med* 2009; 180: 3–10.
28. de Marco R, Accordini S, Cerveri I, *et al.* Incidence of chronic obstructive pulmonary disease in a cohort of young adults according to the presence of chronic cough and phlegm. *Am J Respir Crit Care Med* 2007; 175: 32–39.
29. Scanlon PD, Connett JE, Waller LA, *et al.* Smoking cessation and lung function in mild-to-moderate chronic obstructive pulmonary disease. The Lung Health Study. *Am J Respir Crit Care Med* 2000; 161: 381–390.
30. Anthonisen NR, Connett JE, Murray RP. Smoking and lung function of Lung Health Study participants after 11 years. *Am J Respir Crit Care Med* 2002; 166: 675–679.
31. Vestbo J, Edwards LD, Scanlon PD, *et al.* Changes in forced expiratory volume in 1 second over time in COPD. *N Engl J Med* 2011; 365: 1184–1192.
32. Global Initiative for Chronic Obstructive Lung Disease (GOLD). Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. 2007. www.goldcopd.org Date last updated: February 2013.
33. Anthonisen NR, Connett JE, Kiley JP, *et al.* Effects of smoking intervention and the use of an inhaled anticholinergic bronchodilator on the rate of decline of FEV1. The Lung Health Study. *JAMA* 1994; 272: 1497–1505.
34. Lung Health Study Research Group. Effect of inhaled triamcinolone on the decline in pulmonary function in chronic obstructive pulmonary disease. *N Engl J Med* 2000; 343: 1902–1909.
35. Pauwels RA, Lofdahl CG, Laitinen LA, *et al.* Long-term treatment with inhaled budesonide in persons with mild chronic obstructive pulmonary disease who continue smoking. European Respiratory Society Study on Chronic Obstructive Pulmonary Disease. *N Engl J Med* 1999; 340: 1948–1953.
36. Burge PS, Calverley PM, Jones PW, *et al.* Randomised, double blind, placebo controlled study of fluticasone propionate in patients with moderate to severe chronic obstructive pulmonary disease: the ISOLDE trial. *BMJ* 2000; 320: 1297–1303.
37. Decramer M, Rutten-van Molken M, Dekhuijzen PN, *et al.* Effects of *N*-acetylcysteine on outcomes in chronic obstructive pulmonary disease (Bronchitis Randomized on NAC Cost-Utility Study, BRONCUS): a randomised placebo-controlled trial. *Lancet* 2005; 365: 1552–1560.
38. Soriano JB, Sin DD, Zhang X, *et al.* A pooled analysis of FEV1 decline in COPD patients randomized to inhaled corticosteroids or placebo. *Chest* 2007; 131: 682–689.
39. Tashkin DP, Celli B, Senn S, *et al.* A 4-year trial of tiotropium in chronic obstructive pulmonary disease. *N Engl J Med* 2008; 359: 1543–1554.
40. Decramer M, Celli B, Kesten S, *et al.* Effect of tiotropium on outcomes in patients with moderate chronic obstructive pulmonary disease (UPLIFT): a prespecified subgroup analysis of a randomised controlled trial. *Lancet* 2009; 374: 1171–1178.
41. Celli BR, Thomas NE, Anderson JA, *et al.* Effect of pharmacotherapy on rate of decline of lung function in chronic obstructive pulmonary disease: results from the TORCH study. *Am J Respir Crit Care Med* 2008; 178: 332–338.
42. Suissa S. Lung function decline in COPD trials: bias from regression to the mean. *Eur Respir J* 2008; 32: 829–831.
43. Miller MR, Hankinson J, Brusasco V, *et al.* Standardisation of spirometry. *Eur Respir J* 2005; 26: 319–338.
44. Lung function testing: selection of reference values and interpretative strategies. American Thoracic Society. *Am Rev Respir Dis* 1991; 144: 1202–1218.

45. Quanjer PH, Tammeling GJ, Cotes JE, *et al.* Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. *Eur Respir J* 1993; 6: Suppl. 16, 5–40.
46. Pellegrino R, Viegi G, Brusasco V, *et al.* Interpretative strategies for lung function tests. *Eur Respir J* 2005; 26: 948–968.
47. Fabbri LM, Calverley PM, Izquierdo-Alonso JL, *et al.* Roflumilast in moderate-to-severe chronic obstructive pulmonary disease treated with long-acting bronchodilators: two randomised clinical trials. *Lancet* 2009; 374: 695–703.
48. Rabe KF, Bateman ED, O'Donnell D, *et al.* Roflumilast – an oral anti-inflammatory treatment for chronic obstructive pulmonary disease: a randomised controlled trial. *Lancet* 2005; 366: 563–571.
49. Calverley PM, Rabe KF, Goehring UM, *et al.* Roflumilast in symptomatic chronic obstructive pulmonary disease: two randomised clinical trials. *Lancet* 2009; 374: 685–694.
50. Calverley PMA, MacNee W, Pride NB, *et al.*, eds. Chronic Obstructive Pulmonary Disease. 2nd Edn. London, Arnold, 2003.
51. Ludwig PW, Schwartz BA, Hoidal JR, *et al.* Cigarette smoking causes accumulation of polymorphonuclear leukocytes in alveolar septum. *Am Rev Respir Dis* 1985; 131: 828–830.
52. Bosken CH, Hards J, Gatter K, *et al.* Characterization of the inflammatory reaction in the peripheral airways of cigarette smokers using immunocytochemistry. *Am Rev Respir Dis* 1992; 145: 911–917.
53. Eidelman D, Saetta MP, Ghezzo H, *et al.* Cellularity of the alveolar walls in smokers and its relation to alveolar destruction. Functional implications. *Am Rev Respir Dis* 1990; 141: 1547–1552.
54. Finkelstein R, Fraser RS, Ghezzo H, *et al.* Alveolar inflammation and its relation to emphysema in smokers. *Am J Respir Crit Care Med* 1995; 152: 1666–1672.
55. Saetta M, Di Stefano A, Turato G, *et al.* CD8+ T-lymphocytes in peripheral airways of smokers with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998; 157: 822–826.
56. O'Shaughnessy TC, Ansari TW, Barnes NC, *et al.* Inflammation in bronchial biopsies of subjects with chronic bronchitis: inverse relationship of CD8+ T lymphocytes with FEV1. *Am J Respir Crit Care Med* 1997; 155: 852–857.
57. Singh D, Edwards L, Tal-Singer R, *et al.* Sputum neutrophils as a biomarker in COPD: findings from the ECLIPSE study. *Respir Res* 2010; 11: 77.
58. Gizycki MJ, Hattotuwa KL, Barnes N, *et al.* Effects of fluticasone propionate on inflammatory cells in COPD: an ultrastructural examination of endobronchial biopsy tissue. *Thorax* 2002; 57: 799–803.
59. Hattotuwa KL, Gizycki MJ, Ansari TW, *et al.* The effects of inhaled fluticasone on airway inflammation in chronic obstructive pulmonary disease: a double-blind, placebo-controlled biopsy study. *Am J Respir Crit Care Med* 2002; 165: 1592–1596.
60. Reid DW, Wen Y, Johns DP, *et al.* Bronchodilator reversibility, airway eosinophilia and anti-inflammatory effects of inhaled fluticasone in COPD are not related. *Respirology* 2008; 13: 799–809.
61. Bourbeau J, Christodoulopoulos P, Maltais F, *et al.* Effect of salmeterol/fluticasone propionate on airway inflammation in COPD: a randomised controlled trial. *Thorax* 2007; 62: 938–943.
62. Lapperre TS, Snoeck-Stroband JB, Gosman MM, *et al.* Effect of fluticasone with and without salmeterol on pulmonary outcomes in chronic obstructive pulmonary disease: a randomized trial. *Ann Intern Med* 2009; 151: 517–527.
63. Barnes NC, Qiu YS, Pavord ID, *et al.* Anti-inflammatory effects of salmeterol/fluticasone propionate in chronic obstructive lung disease. *Am J Respir Crit Care Med* 2006; 173: 736–743.
64. Yang IA, Clarke MS, Sim EH, *et al.* Inhaled corticosteroids for stable chronic obstructive pulmonary disease. *Cochrane Database Syst Rev* 2012; 7: CD002991.
65. Rennard SI, Fogarty C, Kelsen S, *et al.* The safety and efficacy of infliximab in moderate to severe chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2007; 175: 926–934.
66. Dentener MA, Creutzberg EC, Pennings HJ, *et al.* Effect of infliximab on local and systemic inflammation in chronic obstructive pulmonary disease: a pilot study. *Respiration* 2008; 76: 275–282.
67. van der Vaart H, Koeter GH, Postma DS, *et al.* First study of infliximab treatment in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2005; 172: 465–469.
68. Kinsella M, Muller NL, Staples C, *et al.* Hyperinflation in asthma and emphysema. Assessment by pulmonary function testing and computed tomography. *Chest* 1988; 94: 286–289.
69. Morrison NJ, Abboud RT, Ramadan F, *et al.* Comparison of single breath carbon monoxide diffusing capacity and pressure–volume curves in detecting emphysema. *Am Rev Respir Dis* 1989; 139: 1179–1187.
70. Dirksen A, Dijkman JH, Madsen F, *et al.* A randomized clinical trial of α_1-antitrypsin augmentation therapy. *Am J Respir Crit Care Med* 1999; 160: 1468–1472.
71. Stolk J, Putter H, Bakker EM, *et al.* Progression parameters for emphysema: a clinical investigation. *Respir Med* 2007; 101: 1924–1930.
72. Dirksen A, Piitulainen E, Parr DG, *et al.* Exploring the role of CT densitometry: a randomised study of augmentation therapy in α_1-antitrypsin deficiency. *Eur Respir J* 2009; 33: 1345–1353.
73. Coxson HO, Dirksen A, Edwards LD, *et al.* The presence and progression of emphysema in COPD as determined by CT scanning and biomarker expression: a prospective analysis from the ECLIPSE study. *Lancet Respir Med* 2013; 1: 129–136.

74. Seemungal TA, Donaldson GC, Paul EA, *et al.* Effect of exacerbation on quality of life in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998; 157: 1418–1422.
75. Miravitlles M, Ferrer M, Pont A, *et al.* Effect of exacerbations on quality of life in patients with chronic obstructive pulmonary disease: a 2 year follow up study. *Thorax* 2004; 59: 387–395.
76. Doll H, Miravitlles M. Health-related QOL in acute exacerbations of chronic bronchitis and chronic obstructive pulmonary disease: a review of the literature. *Pharmacoeconomics* 2005; 23: 345–363.
77. Schmier JK, Halpern MT, Higashi MK, *et al.* The quality of life impact of acute exacerbations of chronic bronchitis (AECB): a literature review. *Qual Life Res* 2005; 14: 329–347.
78. Burrows B, Earle RH. Course and prognosis of chronic obstructive lung disease. A prospective study of 200 patients. *N Engl J Med* 1969; 280: 397–404.
79. Anthonisen NR, Manfreda J, Warren CP, *et al.* Antibiotic therapy in exacerbations of chronic obstructive pulmonary disease. *Ann Intern Med* 1987; 106: 196–204.
80. Rodriguez-Roisin R. Toward a consensus definition for COPD exacerbations. *Chest* 2000; 117: Suppl. 2, 398S–401S.
81. Casaburi R, Mahler DA, Jones PW, *et al.* A long-term evaluation of once-daily inhaled tiotropium in chronic obstructive pulmonary disease. *Eur Respir J* 2002; 19: 217–224.
82. Szafranski W, Cukier A, Ramirez A, *et al.* Efficacy and safety of budesonide/formoterol in the management of chronic obstructive pulmonary disease. *Eur Respir J* 2003; 21: 74–81.
83. Vestbo J, TORCH Study Group. The TORCH (Towards a Revolution in COPD Health) survival study protocol. *Eur Respir J* 2004; 24: 206–210.
84. Donaldson GC, Seemungal TA, Patel IS, *et al.* Longitudinal changes in the nature, severity and frequency of COPD exacerbations. *Eur Respir J* 2003; 22: 931–936.
85. Jones PW, Willits LR, Burge PS, *et al.* Disease severity and the effect of fluticasone propionate on chronic obstructive pulmonary disease exacerbations. *Eur Respir J* 2003; 21: 68–73.
86. Hurst JR. The rhythm of chronic obstructive pulmonary disease exacerbations. *J Clin Epidemiol* 2010; 63: 1285–1286.
87. Saey D, Michaud A, Couillard A, *et al.* Contractile fatigue, muscle morphometry, and blood lactate in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2005; 171: 1109–1115.
88. O'Donnell DE, Bertley JC, Chau LK, *et al.* Qualitative aspects of exertional breathlessness in chronic airflow limitation: pathophysiologic mechanisms. *Am J Respir Crit Care Med* 1997; 155: 109–115.
89. O'Donnell DE, Revill SM, Webb KA. Dynamic hyperinflation and exercise intolerance in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2001; 164: 770–777.
90. O'Donnell DE, Fluge T, Gerken F, *et al.* Effects of tiotropium on lung hyperinflation, dyspnoea and exercise tolerance in COPD. *Eur Respir J* 2004; 23: 832–840.
91. Maltais F, Hamilton A, Marciniuk D, *et al.* Improvements in symptom-limited exercise performance over 8 h with once-daily tiotropium in patients with COPD. *Chest* 2005; 128: 1168–1178.
92. O'Donnell DE, Webb KA. The major limitation to exercise performance in COPD is dynamic hyperinflation. *J Appl Physiol* 2008; 105: 753–755.
93. Belman MJ, Botnick WC, Shin JW. Inhaled bronchodilators reduce dynamic hyperinflation during exercise in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1996; 153: 967–975.
94. O'Donnell DE, Hamilton AL, Webb KA. Sensory-mechanical relationships during high-intensity, constant-work-rate exercise in COPD. *J Appl Physiol* 2006; 101: 1025–1035.
95. O'Donnell DE, Lam M, Webb KA. Measurement of symptoms, lung hyperinflation, and endurance during exercise in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998; 158: 1557–1565.
96. Peters MM, Webb KA, O'Donnell DE. Combined physiological effects of bronchodilators and hyperoxia on exertional dyspnoea in normoxic COPD. *Thorax* 2006; 61: 559–567.
97. Laghi F, Jubran A, Topeli A, *et al.* Effect of lung volume reduction surgery on neuromechanical coupling of the diaphragm. *Am J Respir Crit Care Med* 1998; 157: 475–483.
98. O'Donnell DE, Webb KA, Bertley JC, *et al.* Mechanisms of relief of exertional breathlessness following unilateral bullectomy and lung volume reduction surgery in emphysema. *Chest* 1996; 110: 18–27.
99. Eves ND, Petersen SR, Haykowsky MJ, *et al.* Helium-hyperoxia, exercise, and respiratory mechanics in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2006; 174: 763–771.
100. Porszasz J, Emtner M, Goto S, *et al.* Exercise training decreases ventilatory requirements and exercise-induced hyperinflation at submaximal intensities in patients with COPD. *Chest* 2005; 128: 2025–2034.
101. McGavin CR, Gupta SP, McHardy GJ. Twelve-minute walking test for assessing disability in chronic bronchitis. *Br Med J* 1976; 1: 822–823.
102. Butland RJ, Pang J, Gross ER, *et al.* Two-, six-, and 12-minute walking tests in respiratory disease. *Br Med J (Clin Res Ed)* 1982; 284: 1607–1608.
103. Guyatt GH, Sullivan MJ, Thompson PJ, *et al.* The 6-minute walk: a new measure of exercise capacity in patients with chronic heart failure. *Can Med Assoc J* 1985; 132: 919–923.
104. Knox AJ, Morrison JF, Muers MF. Reproducibility of walking test results in chronic obstructive airways disease. *Thorax* 1988; 43: 388–392.
105. Gerardi DA, Lovett L, Benoit-Connors ML, *et al.* Variables related to increased mortality following out-patient pulmonary rehabilitation. *Eur Respir J* 1996; 9: 431–435.

106. Jones PW. Health status measurement in chronic obstructive pulmonary disease. *Thorax* 2001; 56: 880–887.
107. Spruit MA, Polkey MI, Celli B, *et al.* Predicting outcomes from 6-minute walk distance in chronic obstructive pulmonary disease. *J Am Med Dir Assoc* 2012; 13: 291–297.
108. Carter R, Holiday DB, Nwasuruba C, *et al.* 6-minute walk work for assessment of functional capacity in patients with COPD. *Chest* 2003; 123: 1408–1415.
109. Ischaki E, Papatheodorou G, Gaki E, *et al.* Body mass and fat-free mass indices in COPD: relation with variables expressing disease severity. *Chest* 2007; 132: 164–169.
110. Huijsmans RJ, de Haan A, ten Hacken NN, *et al.* The clinical utility of the GOLD classification of COPD disease severity in pulmonary rehabilitation. *Respir Med* 2008; 102: 162–171.
111. Lee YK, Oh YM, Lee JH, *et al.* Quantitative assessment of emphysema, air trapping, and airway thickening on computed tomography. *Lung* 2008; 186: 157–165.
112. Carr SJ, Goldstein RS, Brooks D. Acute exacerbations of COPD in subjects completing pulmonary rehabilitation. *Chest* 2007; 132: 127–134.
113. Spruit MA, Pennings HJ, Janssen PP, *et al.* Extra-pulmonary features in COPD patients entering rehabilitation after stratification for MRC dyspnea grade. *Respir Med* 2007; 101: 2454–2463.
114. Eiser N, West C, Evans S, *et al.* Effects of psychotherapy in moderately severe COPD: a pilot study. *Eur Respir J* 1997; 10: 1581–1584.
115. de Torres JP, Casanova C, Hernandez C, *et al.* Gender associated differences in determinants of quality of life in patients with COPD: a case series study. *Health Qual Life Outcomes* 2006; 4: 72.
116. Spruit MA, Watkins ML, Edwards LD, *et al.* Determinants of poor 6-min walking distance in patients with COPD: the ECLIPSE cohort. *Respir Med* 2010; 104: 849–857.
117. Swiston JR, Johnson SR, Granton JT. Factors that prognosticate mortality in idiopathic pulmonary arterial hypertension: a systematic review of the literature. *Respir Med* 2010; 104: 1588–1607.
118. Cote CG, Pinto-Plata V, Kasprzyk K, *et al.* The 6-min walk distance, peak oxygen uptake, and mortality in COPD. *Chest* 2007; 132: 1778–1785.
119. Geddes D, Davies M, Koyama H, *et al.* Effect of lung-volume-reduction surgery in patients with severe emphysema. *N Engl J Med* 2000; 343: 239–245.
120. Young J, Fry-Smith A, Hyde C. Lung volume reduction surgery (LVRS) for chronic obstructive pulmonary disease (COPD) with underlying severe emphysema. *Thorax* 1999; 54: 779–789.
121. Griffiths TL, Burr ML, Campbell IA, *et al.* Results at 1 year of outpatient multidisciplinary pulmonary rehabilitation: a randomised controlled trial. *Lancet* 2000; 355: 362–368.
122. Leach RM, Davidson AC, Chinn S, *et al.* Portable liquid oxygen and exercise ability in severe respiratory disability. *Thorax* 1992; 47: 781–789.
123. Corris PA, Neville E, Nariman S, *et al.* Dose–response study of inhaled salbutamol powder in chronic airflow obstruction. *Thorax* 1983; 38: 292–296.
124. Polkey MI, Spruit MA, Edwards LD, *et al.* Six-minute-walk test in chronic obstructive pulmonary disease: minimal clinically important difference for death or hospitalization. *Am J Respir Crit Care Med* 2013; 187: 382–386.
125. Puhan MA, Mador MJ, Held U, *et al.* Interpretation of treatment changes in 6-minute walk distance in patients with COPD. *Eur Respir J* 2008; 32: 637–643.
126. Onorati P, Antonucci R, Valli G, *et al.* Non-invasive evaluation of gas exchange during a shuttle walking test *vs* a 6-min walking test to assess exercise tolerance in COPD patients. *Eur J Appl Physiol* 2003; 89: 331–336.
127. Man WD, Polkey MI, Donaldson N, *et al.* Community pulmonary rehabilitation after hospitalisation for acute exacerbations of chronic obstructive pulmonary disease: randomised controlled study. *BMJ* 2004; 329: 1209.
128. Fuld JP, Kilduff LP, Neder JA, *et al.* Creatine supplementation during pulmonary rehabilitation in chronic obstructive pulmonary disease. *Thorax* 2005; 60: 531–537.
129. Steiner MC, Barton RL, Singh SJ, *et al.* Nutritional enhancement of exercise performance in chronic obstructive pulmonary disease: a randomised controlled trial. *Thorax* 2003; 58: 745–751.
130. Eiser N, Willsher D, Dore CJ. Reliability, repeatability and sensitivity to change of externally and self-paced walking tests in COPD patients. *Respir Med* 2003; 97: 407–414.
131. Das P, Horton R. Rethinking our approach to physical activity. *Lancet* 2012; 380: 189–190.
132. Lee IM, Shiroma EJ, Lobelo F, *et al.* Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet* 2012; 380: 219–229.
133. Parshall MB, Schwartzstein RM, Adams L, *et al.* An official American Thoracic Society statement: update on the mechanisms, assessment, and management of dyspnea. *Am J Respir Crit Care Med* 2012; 185: 435–452.
134. Dyspnea. Mechanisms, assessment, and management: a consensus statement. American Thoracic Society. *Am J Respir Crit Care Med* 1999; 159: 321–340.
135. Fletcher CM, Elmes PC, Fairbairn AS, *et al.* The significance of respiratory symptoms and the diagnosis of chronic bronchitis in a working population. *Br Med J* 1959; 2: 257–266.
136. Bestall JC, Paul EA, Garrod R, *et al.* Usefulness of the Medical Research Council (MRC) dyspnoea scale as a measure of disability in patients with chronic obstructive pulmonary disease. *Thorax* 1999; 54: 581–586.
137. Mahler D. Measurement of dyspnea: clinical ratings. *In*: Mahler D, ed. Dyspnea: Mechanisms, Measurement and Management. New York, Taylor and Francis Inc., 2005; pp. 147–164.

138. Cazzola M, MacNee W, Martinez FJ, *et al.* Outcomes for COPD pharmacological trials: from lung function to biomarkers. *Eur Respir J* 2008; 31: 416–469.
139. Mahler DA, Weinberg DH, Wells CK, *et al.* The measurement of dyspnea. Contents, interobserver agreement, and physiologic correlates of two new clinical indexes. *Chest* 1984; 85: 751–758.
140. Mahler DA, Tomlinson D, Olmstead EM, *et al.* Changes in dyspnea, health status, and lung function in chronic airway disease. *Am J Respir Crit Care Med* 1995; 151: 61–65.
141. Witek TJ Jr, Mahler DA. Minimal important difference of the transition dyspnoea index in a multinational clinical trial. *Eur Respir J* 2003; 21: 267–272.
142. Donohue JF, van Noord JA, Bateman ED, *et al.* A 6-month, placebo-controlled study comparing lung function and health status changes in COPD patients treated with tiotropium or salmeterol. *Chest* 2002; 122: 47–55.
143. Donohue JF, Kalberg C, Emmett A, *et al.* A short-term comparison of fluticasone propionate/salmeterol with ipratropium bromide/albuterol for the treatment of COPD. *Treat Respir Med* 2004; 3: 173–181.
144. Vincken W, van Noord JA, Greefhorst AP, *et al.* Improved health outcomes in patients with COPD during 1 yr's treatment with tiotropium. *Eur Respir J* 2002; 19: 209–216.
145. ZuWallack RL, Mahler DA, Reilly D, *et al.* Salmeterol plus theophylline combination therapy in the treatment of COPD. *Chest* 2001; 119: 1661–1670.
146. Brusasco V, Hodder R, Miravitlles M, *et al.* Health outcomes following treatment for six months with once daily tiotropium compared with twice daily salmeterol in patients with COPD. *Thorax* 2003; 58: 399–404.
147. Mahler D. Dyspnea. *In*: Celli B, ed. Pharmacotherapy of COPD. New York, Marcel Dekker Inc., 2003; pp. 145–158.
148. Aaron SD, Vandemheen KL, Hebert P, *et al.* Outpatient oral prednisone after emergency treatment of chronic obstructive pulmonary disease. *N Engl J Med* 2003; 348: 2618–2625.
149. Mahler DA, Wire P, Horstman D, *et al.* Effectiveness of fluticasone propionate and salmeterol combination delivered *via* the Diskus device in the treatment of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2002; 166: 1084–1091.
150. Hanania NA, Darken P, Horstman D, *et al.* The efficacy and safety of fluticasone propionate (250 μg)/salmeterol (50 μg) combined in the Diskus inhaler for the treatment of COPD. *Chest* 2003; 124: 834–843.
151. Moons P, Budts W, De Geest S. Critique on the conceptualisation of quality of life: a review and evaluation of different conceptual approaches. *Int J Nurs Stud* 2006; 43: 891–901.
152. Carr AJ, Gibson B, Robinson PG. Measuring quality of life: is quality of life determined by expectations or experience? *BMJ* 2001; 322: 1240–1243.
153. Rosenberg R. Health-related quality of life between naturalism and hermeneutics. *Soc Sci Med* 1995; 41: 1411–1415.
154. Jones PW. Issues concerning health-related quality of life in COPD. *Chest* 1995; 107: 187S–193S.
155. Weldam SW, Schuurmans MJ, Liu R, *et al.* Evaluation of quality of life instruments for use in COPD care and research: a systematic review. *Int J Nurs Stud* 2013; 50: 688–707.
156. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993; 118: 622–629.
157. Jones PW. COPD: functional status, health status and primary care. *Prim Care Respir J* 2011; 20: 227–228.
158. Mokkink LB, Terwee CB, Gibbons E, *et al.* Inter-rater agreement and reliability of the COSMIN (COnsensus-based Standards for the selection of health status Measurement Instruments) checklist. *BMC Med Res Methodol* 2010; 10: 82.
159. Mokkink LB, Terwee CB, Patrick DL, *et al.* The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010; 63: 737–745.
160. Mokkink LB, Terwee CB, Knol DL, *et al.* The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol* 2010; 10: 22.
161. Mokkink LB, Terwee CB, Patrick DL, *et al.* The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010; 19: 539–549.
162. McKenna SP, Meads DM, Doward LC, *et al.* Development and validation of the living with chronic obstructive pulmonary disease questionnaire. *Qual Life Res* 2011; 20: 1043–1052.
163. van der Molen T, Willemse BW, Schokker S, *et al.* Development, validity and responsiveness of the Clinical COPD Questionnaire. *Health Qual Life Outcomes* 2003; 1: 13.
164. Stallberg B, Nokela M, Ehrs PO, *et al.* Validation of the Clinical COPD Questionnaire (CCQ) in primary care. *Health Qual Life Outcomes* 2009; 7: 26.
165. Jones PW, Harding G, Berry P, *et al.* Development and first validation of the COPD Assessment Test. *Eur Respir J* 2009; 34: 648–654.
166. Jones PW, Brusselle G, Dal Negro RW, *et al.* Properties of the COPD Assessment Test in a cross-sectional European study. *Eur Respir J* 2011; 38: 29–35.
167. Long term domiciliary oxygen therapy in chronic hypoxic cor pulmonale complicating chronic bronchitis and emphysema. Report of the Medical Research Council Working Party. *Lancet* 1981; 1: 681–686.
168. Continuous or nocturnal oxygen therapy in hypoxemic chronic obstructive lung disease: a clinical trial. Nocturnal Oxygen Therapy Trial Group. *Ann Intern Med* 1980; 93: 391–398.

169. Fishman A, Martinez F, Naunheim K, *et al.* A randomized trial comparing lung-volume-reduction surgery with medical therapy for severe emphysema. *N Engl J Med* 2003; 348: 2059–2073.
170. Rabe KF. Treating COPD – the TORCH trial, P values, and the Dodo. *N Engl J Med* 2007; 356: 851–854.
171. Vestbo J, Anderson J, Brook RD, *et al.* The Study to Understand Mortality and Morbidity in COPD (SUMMIT) study protocol. *Eur Respir J* 2013; 41: 1017–1022.
172. Remington RD. Who should code cause of death in a clinical trial? *Control Clin Trials* 1984; 5: 241–244.

Chapter 2

Asthma

Julia Koopmans and Christian Taube

SUMMARY: Asthma is a heterogeneous syndrome with different phenotypes. This makes it necessary to measure different outcome parameters in clinical trials to adequately assess the response of a patient to a therapy. One issue was the heterogeneous assessment of different outcome parameters, which sometimes made it impossible to compare the results from different studies. Therefore, initiatives by the American Thoracic Society and the European Respiratory Society, as well as by the National Institutes of Health, resulted in recommendations on the assessment of different outcome parameters in asthma. These include symptoms, lung function, exacerbations, health-related quality of life and biomarkers. These outcome parameters describe different dimensions of the disease and a certain panel of parameters has been recommended as a basic outcome parameter for clinical trials in patients with asthma.

Dept of Pulmonology, Leiden University Medical Center, Leiden, The Netherlands.

Correspondence: C. Taube, Dept of Pulmonology, Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, The Netherlands.
Email: c.taube@lumc.nl

Eur Respir Monogr 2013; 62: 19–36.
Copyright ERS 2013.
DOI: 10.1183/1025448x.10044312
Print ISBN: 978-1-84984-044-6
Online ISBN: 978-1-84984-045-3
Print ISSN: 1025-448x
Online ISSN: 2075-6674

Asthma is a heterogeneous syndrome with many different clinical and immunological phenotypes [1, 2]. This is partly due to the broad criteria defining the disease, which results in millions of patients who congregate under this definition. Probably the most accepted definition of asthma is by the Global Initiative for Asthma (GINA), which describes asthma as a chronic inflammatory disorder of the airways. Chronic bronchial inflammation is associated with airway hyperresponsiveness, which leads to the classical symptoms of asthma: recurrent episodes of wheezing, breathlessness, chest tightness and coughing [3]. Today, physicians treating patients with asthma should orient themselves with the available guidelines (*e.g.* the GINA guidelines) to optimise medical therapy. Recommendations in guidelines are based on clinical trials, which sometimes use different outcome parameters. Generally, clinical trials in asthma testing novel therapeutic approaches can be separated into pharmacodynamic, pharmacokinetic, dose–response relationship and the main clinical efficacy studies. The study design and choice of end-points in a study depend on the type of substance under investigation. There are distinctions between reliever or controller substances, which determine outcome parameters as well as the duration of the trials.

One issue when summarising and comparing clinical trials is that different techniques are used to measure certain outcome parameters. Furthermore, relevant outcome parameters have sometimes not been standardised, making the comparison of trials or meta-analysis complicated or not possible [4]. Therefore, different initiatives have been developed to standardise outcome parameters in clinical asthma trials. These include a joint recommendation of the American Thoracic Society (ATS) and European Respiratory Society (ERS) [5] and detailed recommendations based on a workshop lead by several institutes of National Institutes of Health (NIH) [6]. In the present chapter we summarise several outcome parameters for clinical trials in adult patients with asthma, including symptoms, quality of life, lung function and exacerbations, as well as biomarkers.

Clinical outcome parameters

Symptoms

Classical asthma symptoms are shortness of breath, wheezing, night-time awakening and cough. These symptoms directly affect patients and are the reason that patients seek medical treatment. Therefore, the occurrence of symptoms is an important or even the most important outcome parameter in clinical trials. Currently available medical therapy for asthma is not capable of curing patients but is aimed at reduction of asthma symptoms and exacerbations. As a consequence, monitoring of symptoms is mandatory as no surrogate parameters are available. Indeed, correlations between symptoms and other parameters, for example lung function and inflammation, are poor [7–9]. Symptoms are classical patient-reported outcomes. One important problem is that the assessment of symptoms is subjective; strong placebo effects on symptoms like dyspnoea have been identified, making an assessment of untreated response essential in evaluating this outcome parameter [10]. When assessing symptoms several different strategies can be used. One topic for discussion is whether each symptom of asthma (*e.g.* cough and shortness of breath) should be measured and reported individually, or if combined scores can and should be used. It is also necessary to distinguish between the recording of symptoms and the assessment of asthma control. Assessment of current asthma control is a somewhat broader concept, which includes certain clinical manifestations of the disease (symptoms, night-time awakening, reliever use, activity limitation and lung function) [3] that can be measured using different composite scores. Part of the concept of asthma control is the assessment of future risks, including exacerbations, accelerated decline in lung function and side-effects of treatment. It is of interest that nearly all of the instruments described below are affected by copyright, which needs to be taken into account when incorporating these tools into a clinical trial.

Daily diaries

One potential way to record symptoms in patients with asthma is the use of questionnaires in the form of daily diaries. Indeed, several different diaries have been developed for children as well as for adult patients. Some of these diaries are retrospective questionnaires, whereas others are used prospectively to record symptoms during the patients normal day-to-day life. Symptoms are highly variable and it has been shown that a prospective symptom diary is more effective in identifying persistent symptoms in comparison with retrospective symptom recall [11]. In contrast, short-term diaries completed just before study visits do not seem to provide more information compared with a retrospective questionnaire [12]. Furthermore, adherence to daily measurements has been identified as an issue, because adherence to these measurements can be low [13]. Novel electronic devices, also incorporating electronic diaries, can increase adherence. However, mainly due to costs, these devices are not recommended for routine use [14]. Few asthma symptom instruments have been tested for validity, including retrospective symptom questionnaires [15], as well as different asthma diaries for children and adults [16–18]. All these instruments of symptom assessment have been summarised and evaluated recently [14]. Interestingly, data suggest that certain composite scores are comparable to these diaries in their discriminative and evaluative measurement properties [12], making it debatable whether diaries always need to be included in clinical studies or if less cumbersome composite scores of asthma control should be used.

Composite scores and asthma control

International guidelines on asthma management indicate that the primary goal of treatment should be optimum asthma control [3] and, therefore, measures of asthma control are frequently incorporated into clinical studies. The term ''asthma control'' describes the magnitude of the manifestations of asthma and two dimensions have been defined, namely ''current clinical control'' (*e.g.* symptoms, reliever use and lung function) and ''future risk'' (*e.g.* exacerbations and

lung function decline) [9]. Studies incorporating control in their design have shown that the goal of guideline-derived asthma control can be achieved in a majority of patients [19]. Furthermore, it has been confirmed that assessment of current control is also a marker for future asthma control, *e.g.* the risk of developing an exacerbation [20]. However, control needs to be assessed separately from exacerbations, as studies sometimes show differential effects of an intervention on asthma control and asthma exacerbations [21].

Many different options and tools exist to assess asthma control and they have recently been extensively reviewed and evaluated [5, 22]. Indeed, parameters such as episode free days [23–25], symptom days per week [26], asthma control days [27, 28], maximum symptom days [27, 29] and asthma control [30, 31] have been used as outcomes in different clinical studies. Generally, it is necessary to distinguish between semi-quantitative descriptors (*e.g.* total, good or poor control) and instruments in which several independent variables are scored numerically and a score is derived. In clinical studies, numerical variables have been found to be more responsive to change [32]. Another drawback of semi-quantitative assessments is the use of arbitrary cut-off points, which makes the evaluation of these scores difficult. Because of these issues the ATS/ERS recommendations [5], as well as the NIH workshop summary [22], suggest the use of asthma control composite instruments which are expressed as a numerical variables (table 1).

Of the available tools to assess asthma control, the Asthma Control Test (ACT) [31] and Asthma Control Questionnaire (ACQ) [30] are the most frequently used. Both assess symptoms retrospectively, either over 4 weeks (ACT) or over 1 week (ACQ). Both have been validated and cut-off values for uncontrolled asthma (ACT $\leqslant$19 and ACQ $\geqslant$1.5) as well as the minimal clinically important differences have been defined (ACT: 3 and ACQ: 0.5) [31, 33–35]. Both scores correlate well with each other [36]. In contrast to the ACT, the ACQ assessment requires measurement of spirometry, which makes it slightly more complicated to assess in clinical practice. Both of the tests seem to be accurate in distinguishing controlled and poorly controlled asthma. In contrast, for the assessment of uncontrolled asthma the accuracy of ACT has been found to be poor and cut-off points for the ACQ have not been established [37]. Another drawback of the ACQ and ACT is that both do not take treatment into account and ACT or ACQ values with low-dose inhaled treatment should be interpreted differently to the values with high-dose therapy (*e.g.* oral corticosteroids). Indeed, assessment of treatment in this type of a control score could be a further advancement to the current situation. One novel tool is the Composite Asthma Severity Index, which assesses not only symptoms and rescue medication, as well as controller medication and exacerbations, but provides a tool to integrate symptoms in the context of the underlying therapy [38].

Use of rescue medication

The use of rescue mediation is a sign of inadequately controlled asthma and the administration of rescue medication has been included as a parameter in several asthma control composite instruments. When measuring and quantifying rescue medication usage it is usual for the number of inhalations per day to be recorded. It has to be kept in mind that for some rescue medications

Table 1. Minimal recommended asthma trial outcomes

	Baseline measurements	Outcomes for prospective trials
Symptoms	Composite scores (ACQ or ACT)	Composite scores (ACQ or ACT)
Lung function	Spirometry (pre- and post-bronchodilation)	Spirometry (pre-bronchodilation)
Exacerbations	Events 12 months prior to study	Steroid use for at least 3 days
	Steroid use	Hospital admissions for asthma
	Hospital admissions and ED visits	ED/ICU visits because of asthma
Biomarkers	Total IgE and specific IgE screen	None

ACQ: Asthma Control Questionnaire; ACT: Asthma Control Test; ED: emergency department; ICU: intensive care unit. Reproduced and modified from [6] with permission from the publisher.

two inhalations are recommended, whereas others work with one inhalation. Therefore, the use of the term occasions per day, rather that inhalations or puffs per day, has been recommended [5]. It must also be taken into account that some patients use their rescue medication prophylactically, *e.g.* before exercising to prevent exercise-induced bronchoconstriction. Indeed, studies have shown discrepancies between symptom-free days compared to β-agonist-free days, which could be explained by prophylactic use of the rescue medication [39]. Therefore, the ATS/ERS statement recommends reporting rescue β-agonist use separately to prophylactic use and reporting rescue medication as a proportion of reliever free days, as well as the mean number of occasions per day (averaged over 1 week) [5].

Health-related quality of life

Health-related quality of life (HRQoL) is assessed in nearly every interventional study in chronic obstructive pulmonary disease (COPD) patients. It is also assessed in asthma and specific questionnaires have been developed and are frequently used. Noticeably, the use of this outcome parameter has recently become less common in trials of patients with asthma. This could be due to the increased usage of asthma control composite instruments. Other important parameters, such as patient adherence, level of asthma management skills and exposure to stress, are important clinical factors but are often not assessed in clinical trials [40]. Generally, HRQoL has been described as ''functional effects of an illness and its consequent therapy upon a patient as perceived by a patient'' [5, 41]. Asthma-related quality of life, therefore, refers to the impact this disease has on the patient. It is seen as an important outcome [4, 5]. The handling of HRQoL questionnaires can be somewhat cumbersome, as they are often interview based and sometimes quite long. In the workshop report of the NIH, different instruments used to assess HRQoL in adult patients with asthma are described in great detail [4]. These include the Asthma Quality of Life Questionnaire [42–44], Mini-Asthma Quality of Life Questionnaire [45], Living with Asthma [46, 47], St. George's Respiratory Questionnaire [48] and Asthma Questionaire-20 [49]. None of these questionnaires are recommended as a standard questionnaire by the ATS/ERS guidelines [5] or by the NIH workshop [4]. Rather it is advised to select the applicable tools from these questionnaires, which include the specific domains of interest for the study. Furthermore, if translated versions are used it is necessary to ensure that validation for a translation and back translation of the questionnaire has occurred.

Exacerbations

Asthma exacerbations are a transient worsening of asthma and define events which change the patient's health status. Exacerbations may be triggered by different aetiologies [50]. These includes exposure to allergens, irritants, exercise, air pollutants or viral respiratory tract infections, predominantly those caused by human rhinoviruses [51], or the reduction of inhaled steroid treatment [52]. It is important to distinguish exacerbations from poorly controlled or uncontrolled asthma [53]. Analyses of exacerbations in clinical trials have been performed in patients with mild and more severe disease and different therapeutic strategies to reduce exacerbations have been assessed [54]. In patients with mild asthma the effectiveness of anti-inflammatory treatment when assessed by lung function parameters can be impaired, as lung function is often normal in this patient group. Also, these patients often show few symptoms making the assessment of symptom scores difficult. By contrast, evaluation of asthma exacerbations is often used and indeed treatment with inhaled steroids leads to a reduction of acute exacerbations compared with placebo treatment [26]. In moderate-to-severe asthma patients exacerbations have been analysed to establish the effectiveness of medical treatments, for example combination therapy with inhaled steroids and long acting β_2-agonists [39, 55]. In patients with severe asthma, exacerbations are an important outcome parameter and are frequently used as primary end-points, for example in the assessment of novel biological therapies, such as monoclonal antibodies against either IgE [56–60], interleukin (IL)-5 [61–63], IL-13 [64] or the α-subunit of the IL-4 receptor [65].

One caveat, which has to be considered, is that exacerbations have been described using a wide variety of definitions ranging from episodes occurring several times a week to severe events, which require hospital treatment. Given this broad use of the term a standardisation of the definition is necessary to allow comparison between different clinical trials. Therefore, the ATS/ERS recommendation on definitions of exacerbations in clinical trials was a necessary advancement in standardisation of this outcome parameter [5]. In the recommendations severe asthma exacerbations are defined as "events that require urgent action on the part of the patient and physicians to prevent serious outcome, such as hospitalization or death from asthma" [5]. This mainly includes the use of systemic corticosteroids. Furthermore, is it recommended that all emergency department visits, hospitalisations, stays in intensive care units (ICUs) and intubations are reported [53]. Moderate asthma exacerbations were defined as troublesome events resulting in a temporary change of medications. These episodes do not require the use of systemic steroids, emergency room visits or hospitalisation for asthma [5]. In contrast, mild exacerbations were not differentiated. These episodes range just outside the normal variation of the disease and are, therefore, not clearly distinguishable from transient loss of asthma control [54]. However, this stratification system has been debated and the recommendations of the NIH workshop do not recommend stratification into moderate or severe exacerbation [53].

Mortality

While asthma is extremely common, fatal or near fatal asthma is rare. Therefore, mortality in asthma is usually not assessed as a primary or secondary end-point parameter in asthma clinical trials. The death of a patient in a clinical trial is a severe adverse event and it is recommended that in all asthma clinical trials all-cause mortality and asthma-related mortality is reported [53]. The validity of disease specific mortality as an outcome parameter is very much dependent on the accurate determination of the cause of death. In a large analysis of nearly 95 000 patients with asthma only 10 asthma-related deaths were observed [66]. The low occurrence of deaths in this population makes mortality an impractical end-point for clinical asthma studies. However, for decades there has been discussion over the safety of β_2-agonists in the treatment of asthma. Chronic use of long-acting β_2-agonists has been associated with asthma-related death in a few asthma studies [67, 68]. This was initially linked to treatment of patients with long acting β_2-agonists without concomitant steroid treatment. Still, the question has been put forward as to whether treatment with long-acting β_2-agonists poses a risk for patients even when an inhaled steroid is used on a regular basis [69]. Regulatory authorities in the USA have, therefore, mandated a huge effort to investigate this issue. Four randomised identically designed trials, each involving 11 700 patients and a smaller trial involving 6200 will be performed [70]. Even with this large number of patients a composite end-point of serious asthma outcomes was chosen: asthma-related death, intubation or hospitalisation, rather than mortality alone. This again demonstrates that mortality, due to the low frequency in this patient group, is not a feasible outcome parameter for asthma trials. Exceptions are studies in which treatment strategies for acute life-threatening attacks are evaluated. Indeed, in this type of study admission to an ICU, respiratory failure, need for endotracheal intubation and mechanical ventilation, as well as mortality, are important outcome parameters [71–75].

Lung function

Asthma is an obstructive airway disease and increasing loss of lung function can occur in certain asthma patients. Therefore, measurement of lung function is an important assessment in clinical trials. Lung function needs to be measured at the beginning of trials to adequately characterise the patients included in the trial. Furthermore, lung function assessment and changes during the study are an important trial outcome. In previous decades, lung function was used as the primary or main outcome parameter in many clinical trials. Indeed, airflow limitation is a major cause of dyspnoea and chest tightness, but functional parameters are only weakly correlated with symptoms.

Therefore, other outcome parameters, *e.g.* exacerbations and asthma control, have become more and more important to assess the complexity of this disorder. Many different lung function parameters can be assessed, but we will mainly address spirometry, peak expiratory flow, assessment of lung volumes and measurement of airway hyperresponsiveness.

Spirometry

Spirometry measures how an individual inhales or exhales volumes of air as a function of time and the outcome signals may be volume or flow [76]. Indeed, spirometry is highly standardised [76], which allows the assessment of airway obstruction as an important outcome in asthma. Due to the high level of standardisation it is possible to directly compare results obtained in different studies and it is possible to analyse this data in meta-analyses containing different trials. The main parameters assessed with spirometry are vital capacity (VC), either measured as forced vital capacity (FVC) or inspiratory vital capacity, and forced expiratory volume in 1 s (FEV_1), as well as the FEV_1/FVC ratio. In addition, peak expiratory flow (PEF) can be derived from this measurement. Other values, like the forced expiratory flow at 25–75% of FVC, are more variable and, therefore, of less importance and will not be discussed here.

In terms of evaluating spirometric values several recommendations have been published with regard to reporting lung function values in clinical studies [5, 77]. In many trials VC data are expressed as FVC. At baseline both FVC and FEV_1 should be expressed as per cent predicted. The recommendation of the ATS/ERS spirometry Task Force is to use the European Coal and Steel reference values for Europe [78] and other reference values for North America [79]. However, a Task Force organised by the ERS (ERS Global Lung Function Initiative) has recently published new spirometric prediction equations which can be applied globally to different ethnic groups within an age range from 3 to 95 years [80].

When assessing spirometric values it is necessary to distinguish between pre-bronchodilator and post-bronchodilator measurements. Indeed, pre-bronchodilator values are assessed by withholding bronchodilating drugs (and potentially also the study medication), which includes no administration of short-acting bronchodilators for more than 6 h, long-acting bronchodilators (*e.g.* formoterol, salmeterol or acclidinium) for more than 12 h and very long-acting bronchodilators (*e.g.* indacaterol, tiotropium and glycoporonium) for more than 24 h. Post-bronchodilator values are usually assessed 15 min following the administration of a rapid acting bronchodilator, such as salbutamol (usually at a dose of 400 μg) [5, 77]. With this dose the plateau in lung function response is usually reached and, therefore, it is not mandatory to stop study medication or other bronchodilator therapy if only this outcome is determined. It is recommended to measure pre-bronchodilator and post-bronchodilator lung function values at the start of a trial. Depending on the type of study either pre- or post-bronchodilator values should be assessed and reported as an outcome. For prospective clinical studies pre-bronchodilator values are mandatory (table 1). In contrast, both pre- and post-bronchodilator lung function values should be determined in observational studies [77].

The usefulness of determining bronchodilator reversibility (by assessing pre- and post-bonchodilator FEV_1 values) is a point of debate, as this test has often been used as an inclusion criterion to discriminate between COPD and asthma. It is thought that the response to inhaled bronchodilators reflects the magnitude of airway smooth muscle relaxation and, therefore, should be responsive in patients with asthma. However, it is now very well recognised that patients with COPD can show a response larger than 200 mL and 12% to inhaled bronchodilators [81], and that some patients with asthma show fixed airflow obstruction [82]. It is still recommended to measure and report bronchodilator reversibility as part of the baseline measurements of patients to adequately describe the patient population [5, 77]. Values should be expressed as percentage change from pre- to post-bronchodilator (calculated from the absolute values), as well as the absolute change [77].

Peak expiratory flow

PEF is the maximum expiratory flow achieved from a maximum forced expiration [76] and can be obtained from flow–volume curve data or using a peak flow meter. The effort compared with FEV_1 is usually shorter and it is somewhat easier to perform. Still, as a measured parameter of airway obstruction PEF is thought to be inferior to FEV_1. This perception is due to the fact that PEF is not as stable in terms of reproducibility and may underestimate airway obstruction in individuals with airway remodelling. The biggest advantage of peak flow measurements is that they are easy to perform and it is possible to perform frequent repeat measurements of the patient at home, which allows comparison over time. When assessing PEF it is recommended to report absolute changes over time as well as percentage changes calculated from the absolute values [77]. For daily measurements is it also possible to report the variability of peak flow values. Indeed, recent studies have used sophisticated models to assess the peak flow fluctuation values demonstrating that the characterisation of fluctuations provides a basis for risk prediction in this patient group [83, 84]. While short-term compliance with PEF measurements is fairly good, long-term compliance with measurement of PEF on a daily basis is low and this might limit its usefulness in clinical trials. However, the utilisation of electronic peak flow meters seems to improve the compliance of patients in performing these measurements [85, 86].

Bronchial hyperresponsiveness

Bronchial hyperresponsiveness (BHR) is a key feature of asthma and indeed the absence of BHR in untreated patients rules out the diagnosis of asthma with reasonable certainty. Indeed, proof of BHR is often used in clinical trials as an inclusion criterion to adequately include patients with asthma. The pathogenesis of BHR in asthma is characterised by a variety of triggers and abnormalities of effector structures in the airways, such as smooth muscle cells and nerves [87]. BHR can be measured using direct or indirect challenges. As a direct challenge inhalation of histamine or methacholine is used. Both substances act directly on airway smooth muscle and are most commonly used in challenge tests [88]. By contrast, indirect challenges cause the release of endogenous mediators that cause the airway smooth muscle to contract and the airways to narrow [89]. Commonly used indirect challenges include exercise, eucapnic voluntary hyperpnoea, hypertonic saline, cold air exposure, inhalation of adenosine monophosphate and mannitol [90–93]. Indirect challenges are interesting with regard to pathophysiological mechanisms and the effect of different interventions on indirect challenge tests have been described [89].

However, indirect tests are not that well standardised. For example, eucapnic voluntary hyperpnoea can indeed induce bronchoconstriction in many patients with asthma. However, this test is not considered a diagnostic standard for patients with asthma, due to missing standardisation and the complexity of necessary technology. Of the indirect tests the exercise challenge test is the best standardised and international guidelines exist [94]. These guidelines recommend the use of a motor driven treadmill or cycle ergometer and control of the temperature and humidity of inspired air to reduce response variability. The recommended outcome parameter is per cent reduction in FEV_1 from baseline [77]. Histamine and methacholine challenges are very well standardised. Indeed, an ATS guideline recommends the use of a dosimeter method using a calibrated nebuliser and dosing schedule [94]. However, other protocols have also been described, which deliver results comparable to the dosimeter [95, 96]. For clinical trials the dosimeter method has been recommended [5] and the reported outcome parameter should be the provocative concentration causing a 20% fall in FEV_1 [77]. In many studies BHR is assessed at baseline to describe the study population, but is not often repeatedly measured during the study [97].

Long volumes and airway resistances

Lung hyperinflation is associated with the sensation of dyspnoea and chest tightness [98]. Lung hyperinflation can be found in patients with acute asthma, but in patients with more severe disease

a persistent lung hyperinflation has also been reported [99–102]. One technique to measure hyperinflation is body plethysmography, which allows assessment of long volumes such as functional residual capacity (FRC). In combination with deep expiration and inspiration, total lung capacity (TLC) and residual volume (RV) can also be assessed [103]. A second option for assessment of airway volumes is the use of dilutional gas methods, which also allows for assessment of RV. Measurement of FRC with body plethysmography allows FRC assessment of the whole lung (including parts of the lung which are not ventilated), while helium dilution assesses the FRC of the ventilated parts of the lung. Therefore, comparison of FRC obtained *via* body plethysmography and helium dilution has been described as a useful method for measurement of the overall unventilated lung volume [104]. If lung volumes are measured then FRC, RV and TLC (% predicted), as well as FRC/TLC and RV/TLC (both ratios taken from absolute values) should be reported [5, 77].

Body plethysmography is safe and noninvasive, but requires certain equipment and skilled personnel. Therefore, body plethysmography is not usually performed in larger clinical trials. This is unfortunate because body plethysmography also allows assessment of airway resistance (R_{aw}) by the direct measurement of specific R_{aw} and the subsequent calculation of R_{aw} as the ratio of specific airway resistance to FRC [103]. R_{aw} is less dependent on patient effort compared with forced volumes and provides an assessment of airway calibre [105]. Another technique to assess pulmonary mechanics and R_{aw} is the forced oscillation technique (FOT). FOT is a noninvasive test which provides unique information about lung mechanics that is not available from spirometry or body plethysmography, *e.g.* the mechanical impedance of the respiratory system. FOT employs small amplitude pressure oscillations superimposed onto the normal breathing and, therefore, is independent from performance of different respiratory manoeuvres. Also, these measurements have been standardised [106]. Despite the additional value of this assessment of lung mechanics FOT is not recommended as a standard parameter in clinical asthma trials [5, 77].

Biomarkers and airway inflammation

Asthma is primarily considered an inflammatory disorder. Treatment of asthma is mainly aimed at control of symptoms, although the occurrence of symptoms and lung function are only weakly associated with markers of bronchial inflammation. Moreover, clinically masked increases in bronchial inflammation can occur in guideline treated asthma [107]. Importantly, the underlying inflammatory process, in particular, might be causally related to increased loss of lung function throughout life in certain types of asthma [108–110]. Phenotyping of patients with asthma into groups with shared clinical characteristics and, more importantly, with similarities in the bronchial inflammatory process, has led to the recognition of steroid responsive and less steroid responsive forms of asthma, the latter group comprising up to 50% of all asthmatics [111, 112]. The most important clinical phenotypes come from cluster analyses [111, 113, 114] and an important current categorisation involves two immunologically determined groups, namely T-helper cell (Th) type 2-associated and non-Th2-associated types [111, 115]. The Th2-associated types comprise early onset allergic, late onset eosinophilic and exercise-induced asthma; the non-Th2-associated types cover obesity-related and neutrophilic asthma. Molecular understanding of the inflammatory processes in these phenotypes is increasing. Therefore, different biomarkers of airway inflammation are being evaluated as outcome parameters in clinical trials, these have recently been summarised extensively [116]. Some of these biomarkers play an important role in characterising the study population, predicting the response to an intervention or evaluating the effectiveness of an intervention. Important progress has been made in standardising the measurement of some of these markers [117, 118], as well as recommendations as to which of these biomarkers are considered required or optional in clinical trials in asthma [116]. It needs to be kept in mind that different biomarkers, *e.g.* exhaled nitric oxide fraction (F_{eNO}), IgE levels, blood eosinophil and neutrophil counts, poorly predict the amount of sputum eosinophil and neutrophil percentages [119]. Future studies are needed to determine the usefulness of these biomarkers in assessing the effects of therapeutic interventions on asthma outcome (table 2).

Table 2. Treatment effect of different therapeutic interventions on biomarkers, lung function and bronchial hyperresponsiveness (BHR) in asthma

	*F*eNO	Sputum		Blood eosinophils	FEV1	BHR	[Ref.]
		Eosinophils	Neutrophils				
Corticosteroids	↓ ↓	↓ ↓	↑ ↑	↓ ↓/–	↑ ↑	↓ ↓	[120, 121–123]
Anti-IgE	–	↓ ↓	NA	↓ ↓	↑	–	[56, 60, 124–126]
Anti-IL-5	–	↓ ↓	NA	↓ ↓	–	–	[61–63, 127, 128]
Anti-IL-13	↓ ↓	NA	NA	–	↑	NA	[64, 129]
Anti-IL-4-Rcα	↓ ↓	NA	NA	NA	↑	NA	[65]

*F*eNO: exhaled nitric oxide fraction; FEV1: forced expiratory volume in 1 s; IL: interleukin; IL-4-Rcα: IL-4 receptor α subunit; ↑ ↑: significant increase; ↓ ↓: significant decrease; ↑: significant increase but <10%; –: no significant change; NA: not available.

Exhaled nitric oxide

*F*eNO has been recommended as an outcome parameter in clinical trials in asthma to evaluate the effectiveness of asthma treatments [116]. This parameter plays a role in characterising the study population [111]. The measurement of *F*eNO has been standardised, is easy to perform and is highly reproducible [117]. However, it is important to recognise that the measurement of *F*eNO can be confounded by several factors. Smoking, for instance, decreases *F*eNO levels and contamination from the nose can lead to very high levels [117, 130]. *F*eNO is particularly elevated in Th2-related asthma, although there is overlap with healthy and atopic subjects [118]. Levels of *F*eNO have been linked to airway inflammation [131] and *F*eNO has been postulated as a biomarker of bronchial inflammation in asthma. Importantly, *F*eNO levels lower than 25 ppb in adult asthmatics indicates a low likelihood of responsiveness to corticosteroids [120, 132–134], and there is also evidence that *F*eNO is predictive of a steroid response in non-eosinophilic asthma [120]. Therefore, one of the potencies of *F*eNO could probably be the reduction of unnecessary steroid use. Furthermore, increased levels of *F*eNO have been shown to correlate with the effect of therapeutic interventions such as anti-IgE or IL-13 inhibition [64, 65, 135]. In contrast, treatment with anti-IL-5, which reduces eosinophilic airway inflammation, had no effect on *F*eNO (table 2). These findings show that sputum eosinophils are regulated in a different way to *F*eNO and that both markers may reflect different aspects of the inflammatory process in asthma [120, 132, 134, 136]. There is still some controversy regarding the use of *F*eNO in clinical practice. Clinical trials on inclusion of *F*eNO in adjusting anti-inflammatory treatment of asthma did not show uniform results [136–138].

Induced sputum

The measurement of sputum eosinophils is well validated, but the complexity of the measurement and the success rate of ~70% limit its applicability in the clinical setting. Despite the lack of a gold standard for sputum induction, certain technical and operational factors have been recommended [139]. Like *F*eNO, sputum eosinophils increase after bronchial allergen challenge and decrease after treatment with corticosteroids. Increased numbers of eosinophils in the bronchial airways are traditionally considered a hallmark of asthma [111]. The percentage of sputum eosinophils has been recommended as an outcome in clinical trials in asthma to evaluate the effectiveness of asthma treatments and is also important for characterising the study population [111, 116]. As mentioned previously, different phenotypes with eosinophilic and non-eosinophilic inflammation have been described. The distinction between these forms is mainly based on a sputum eosinophil level above or below 2% [140]. Levels of sputum eosinophils have been broadly correlated with disease severity and are predictive of the response to corticosteroids. IL-5 is considered to be a key

cytokine for the recruitment of eosinophils to the airways, but eosinophils can also be attracted by certain chemokines, like eotaxin or RANTES. In addition to the assessment of eosinophil numbers in sputum, numbers of neutrophils are also measured to identify the neutrophilic asthma phenotype, which is considered a more severe type that is less- or nonresponsive to steroids [140]. There is still some debate whether a distinct neutrophilic phenotype exists or if it would be appropriate to distinguish patients with eosinophilic and non-eosinophilic inflammation [112].

Similar to *F*eNO, differential effects of treatment with molecular-based therapies have been shown with respect to sputum eosinophils (table 2). For instance, a single dose of anti-IL-5 before allergen challenge significantly reduced sputum and blood eosinophils, but had no effect on the late asthmatic response or on airway hyperresponsiveness to histamine [127]. More recently, treatment with anti-IL-5 in more severe eosinophilic forms of asthma showed significant reductions in eosinophils and asthma exacerbations but no change in lung function, quality of life or *F*eNO [63]. *F*eNO levels, IgE levels, blood eosinophil and neutrophil counts, FEV1 and age have been described as poor surrogates for predicting sputum eosinophil numbers [119]. These findings confirm the lack of a strong relationship between these outcomes and markers of airway inflammation, demonstrating the necessity of assessing all these parameters. Importantly, the inclusion of sputum eosinophils in tailoring the adjustment of therapy showed a significant reduction in exacerbations when compared with traditional ways of tailoring asthma treatment, whereas, as yet, this could not be shown for *F*eNO-based tailoring of asthma treatment [141–143]. These findings illustrate the importance of the right choice of outcomes in clinical trials and stress the importance of including sputum eosinophils and exacerbations in asthma trials wherever possible besides the minimal recommended outcomes (table 1), thus hopefully adequately defining a response to an intervention, especially when evaluating novel anti-inflammatory concepts.

Markers in serum

Distinguishing allergic from non-allergic asthma can be important for certain treatment choices. Quantitative serological measures of total IgE and individual allergen-specific IgE antibodies are considered supplemental biomarker outcomes [116]. It is important to realise that 40–50% of the western population is sensitised to one or more aeroallergen [144] and that the prevalence of asthma is only 5–10% [145]. Therefore, a large group of sensitised subjects does not experience any symptoms of allergy and asthma, whereas others develop involvement of other organ systems, like eczema or rhinoconjunctivitis, but not asthma. Thus, in patients with asthma there is not always a causative relationship between sensitisation and asthma symptoms. For instance, in an obese patient with asthma who is sensitised to some extent, phenotyping based on the level of sputum eosinophils may be important in differentiating between obesity-related and allergic asthma. Interestingly, neither serum IgE [146, 147] nor allergen-specific IgE [148] have been particularly predictive of response to anti-IgE therapy. Therefore, a multi-allergen screen is mainly considered a biomarker that permits characterisation of the atopic status of a study population, but is not suggested as an outcome parameter [116].

Another serum marker, which has gained increased attention, is periostin. Periostin is generated by airway epithelial cells in response to IL-13 exposure [149] and can be measured in serum. Recent studies have identified periostin as a marker of Th2-related airway inflammation in patients with asthma [115, 150]. Even more interesting, the level of serum periostin has been associated with a greater improvement in FEV1 following therapy with anti-IL-13 [64], the change in FEV1 was less than 10%. High serum periostin levels at baseline have also been associated with a greater decrease in exacerbations after treatment with anti-IgE; however, the difference compared to low serum periostin was not significant [135]. So far no standardised test for serum periostin is commercially available and, therefore, the inclusion of this parameter in clinical trials is limited at the present.

Translation of outcome parameters into clinical practice

A central goal of clinical trials is the improvement of patient care. Evidence-based medicine is a cornerstone in treatment decision making, but this can only be achieved with valid data from clinical studies. A vigorous debate has emerged on the validity of results from investigator-initiated and industry-sponsored trials. One issue in this debate has been the role of the pharmaceutical industry and potential conflicts of interest, which might arise from this involvement [151]. Needless to say, pharmaceutical companies are dependent on clinical trials to demonstrate the effectiveness and safety of their own compounds. It has been assumed that this conflict of interest could result in a so-called publication bias preventing the publication of unfavourable results [151, 152] and indeed such cases have occurred [153]. In addition, it is noticeable that trials funded by pharmaceutical industry are more likely to have positive outcomes for the investigated drug when compared with non-pharmaceutical funded trials, *e.g.* by choosing favourable, but clinical irrelevant, primary end-points [151, 154]. Some of these patterns have also been observed in clinical asthma trials. Indeed, publications assessing the effectiveness of the combination of long-acting β-agonists with inhaled steroids in comparison to inhaled steroids alone with a first or senior author affiliated with industry were 1.5 times more likely to report statistical significant results for the primary outcome parameter compared with studies with other types of industry involvement [155]. In addition, if the primary outcome parameter is negative, studies funded by industry are more likely to report subgroup analyses than non-industry funded trials [156]. Therefore, when appraising results from clinical trials the involvement of a pharmaceutical company and any potential conflict of interest also needs to be considered.

To minimise potential conflicts of interest and increase transparency in medical publishing several initiatives have been started. One demand is that companies that have sponsored randomised controlled trials (RCTs) have independent statistical reviews performed on their data to ensure unbiased analysis [151]. To reduce the potential for publication bias and also to increase transparency with regard to defined primary and secondary end-points the International Committee of Medical Journal Editors (ICMJE) has issued a statement that, as of 2004, phase III clinical studies will only be considered for publication if these trials have been registered before enrolment [157]. In addition, since 2008 all clinical trials, including phase I and pharmacokinetic studies, need to be registered beforehand [158] and progress must be updated every 6 months during the recruitment phase [159]. Indeed, an increasing number of clinical trial registries are available, which participate in the World Health Organization International Clinical Trial Registry Platform (http://who.int/ictrp/network/primary/en/index.html). To make the potential conflicts of interest of authors more apparent the ICMJE developed an electronic uniform disclosure form and placed it in the public domain in October 2009 [160]. In the future the vast majority of novel products will be developed by pharmaceutical industry and will be tested in clinical trials supported or run by industry. In this context it is necessary that these trials adhere to strict scientific and ethical standards to ensure conflicts of interest are minimised and to provide unbiased scientific information which can be used to improve patient care.

In addition to potential conflicts of interest, it is always necessary to account for several shortcomings of clinical trials in the "real life" situation, which sometimes complicate extrapolation of the clinical study findings to treatment decisions for individual patients. One example is patients with asthma, who actively smoke or have a smoking history of more than 10 pack-years. These patients are excluded in most clinical asthma trials. Further examples for frequent exclusion are comorbidities, limitation of lung function, absence of bronchial reversibility and age. Applying these rules to the "real life" patients seen in a clinical centre resulted in only about 5% of patients with asthma meeting the inclusion criteria [161]. Adding further criteria, like being symptomatic or regular use of inhaled corticosteroids, resulted in a further drop of eligible patients to about 3%. This illustrates that caution is necessary before transferring findings from clinical trials with a highly selective patient group into real life. In addition, patients in RCTs are usually carefully looked after by the study team. Compliance and adherence to study and

concomitant medication is vigorously monitored. By contrast, in real life adherence and compliance to asthma medication are a major issue. Therefore, in recent years so-called real life or pragmatic studies, conducted with open-label therapy, have become increasingly popular [162, 163]. Indeed, a recent pragmatic study showed that treatment with a leukotriene receptor antagonist was equivalent to inhaled corticosteroids as a first-line controller therapy [164]. These findings are in contrast to many randomised, placebo-controlled, double-blind clinical trials, which demonstrated a higher effectiveness of inhaled steroids compared with leukotriene receptor antagonists [165]. This discrepancy might be partly explained by adherence to therapy, which for inhaled steroids is low in real life and somewhat higher for leukotriene receptor antagonists [164, 166]. Overall, these findings show that findings in clinical trials should not be uncritically adapted to every patient. Outcomes in clinical trials can help find the right therapy, but an assessment of the individual patient and appropriate adaptation of therapy always remains necessary.

Conclusion

Asthma is a heterogeneous disease with different phenotypes and, therefore, adequate assessment of patients with different outcome variables is necessary in clinical trials. In recent years there has been a development from assessing primarily physiological parameters (*e.g.* FEV_1) to increasingly measuring symptoms (*e.g.* composite scores) and asthma exacerbation as outcome parameters. Furthermore, novel biomarkers have been implemented which can provide additional insights. This is increasingly important, as more specific approaches like treatments with monoclonal antibodies for certain targets are tested, which seem to be especially effective in certain asthma patients, who can be phenotyped adequately using biomarkers. There have been meaningful initiatives to standardise outcome parameters by the ATS/ERS or the NIH, which have led to several recommendations for the standardisation of clinical trial outcome parameters. This includes recommendations for the adequate assessment of symptoms, lung function, quality of life, exacerbations and biomarkers. It is apparent that these different outcome parameters express different domains of the asthma syndrome and that they cannot be used as surrogate parameters for each other. A core of outcome variables have been defined (table 1), which should be incorporated in clinical asthma trials; however, depending on the type of clinical study and the investigated compound, different outcome parameters can be added. Physicians still need to be cautious and not uncritically transfer findings from clinical trials into the real life situation. Differences between the study population and the actual treated patient need to be considered.

Statement of Interest

C. Taube has received personal fees for speaking from Boehringer Ingelheim, Novartis, Almirall, AstraZeneca and GlaxoSmithKline.

References

1. Wenzel S. Severe asthma: from characteristics to phenotypes to endotypes. *Clin Exp Allergy* 2012; 42: 650–658.
2. Wenzel SE. Asthma: defining of the persistent adult phenotypes. *Lancet* 2006; 368: 804–813.
3. Global Initiative for Asthma. Global strategy for asthma management and prevention: updated December 2012. GINA, 2012. Available from: www.ginasthma.org/uploads/users/files/GINA_Report_2012.pdf
4. Wilson SR, Rand CS, Cabana MD, *et al.* Asthma outcomes: quality of life. *J Allergy Clin Immunol* 2012; 129: Suppl. 3, S88–S123.
5. Reddel HK, Taylor DR, Bateman ED, *et al.* An official American Thoracic Society/European Respiratory Society statement: asthma control and exacerbations: standardizing endpoints for clinical asthma trials and clinical practice. *Am J Respir Crit Care Med* 2009; 180: 59–99.
6. Busse WW, Morgan WJ, Taggart V, *et al.* Asthma outcomes workshop: overview. *J Allergy Clin Immunol* 2012; 129: Suppl. 3, S1–S8.
7. Rosi E, Ronchi MC, Grazzini M, *et al.* Sputum analysis, bronchial hyperresponsiveness, and airway function in asthma: results of a factor analysis. *J Allergy Clin Immunol* 1999; 103: 232–237.
8. Crimi E, Spanevello A, Neri M, *et al.* Dissociation between airway inflammation and airway hyperresponsiveness in allergic asthma. *Am J Respir Crit Care Med* 1998; 157: 4–9.

9. Taylor DR, Bateman ED, Boulet LP, *et al.* A new perspective on concepts of asthma severity and control. *Eur Respir J* 2008; 32: 545–554.
10. Wechsler ME, Kelley JM, Boyd IO, *et al.* Active albuterol or placebo, sham acupuncture, or no intervention in asthma. *N Engl J Med* 2011; 365: 119–126.
11. Reznik M, Sharif I, Ozuah PO. Classifying asthma severity: prospective symptom diary or retrospective symptom recall? *J Adolesc Health* 2005; 36: 537–538.
12. Juniper EF, O'Byrne PM, Ferrie PJ, *et al.* Measuring asthma control. Clinic questionnaire or daily diary? *Am J Respir Crit Care Med* 2000; 162: 1330–1334.
13. Stone AA, Shiffman S, Schwartz JE, *et al.* Patient non-compliance with paper diaries. *BMJ* 2002; 324: 1193–1194.
14. Krishnan JA, Lemanske RF Jr, Canino GJ, *et al.* Asthma outcomes: symptoms. *J Allergy Clin Immunol* 2012; 129: Suppl. 3, S124–S135.
15. Revicki DA, Leidy NK, Brennan-Diemer F, *et al.* Integrating patient preferences into health outcomes assessment: the multiattribute Asthma Symptom Utility Index. *Chest* 1998; 114: 998–1007.
16. Santanello NC, Barber BL, Reiss TF, *et al.* Measurement characteristics of two asthma symptom diary scales for use in clinical trials. *Eur Respir J* 1997; 10: 646–651.
17. Santanello NC, Davies G, Galant SP, *et al.* Validation of an asthma symptom diary for interventional studies. *Arch Dis Child* 1999; 80: 414–420.
18. Santanello NC, Demuro-Mercon C, Davies G, *et al.* Validation of a pediatric asthma caregiver diary. *J Allergy Clin Immunol* 2000; 106: 861–866.
19. Bateman ED, Boushey HA, Bousquet J, *et al.* Can guideline-defined asthma control be achieved? The Gaining Optimal Asthma ControL study. *Am J Respir Crit Care Med* 2004; 170: 836–844.
20. Bateman ED, Reddel HK, Eriksson G, *et al.* Overall asthma control: the relationship between current control and future risk. *J Allergy Clin Immunol* 2010; 125: 600–608.
21. Gibson PG, Powell H, Ducharme FM. Differential effects of maintenance long-acting beta-agonist and inhaled corticosteroid on asthma control and asthma exacerbations. *J Allergy Clin Immunol* 2007; 119: 344–350.
22. Cloutier MM, Schatz M, Castro M, *et al.* Asthma outcomes: composite scores of asthma control. *J Allergy Clin Immunol* 2012; 129: Suppl. 3, S24–S33.
23. Bacharier LB, Phillips BR, Zeiger RS, *et al.* Episodic use of an inhaled corticosteroid or leukotriene receptor antagonist in preschool children with moderate-to-severe intermittent wheezing. *J Allergy Clin Immunol* 2008; 122: 1127–1135.
24. Guilbert TW, Morgan WJ, Zeiger RS, *et al.* Long-term inhaled corticosteroids in preschool children at high risk for asthma. *N Engl J Med* 2006; 354: 1985–1997.
25. Tasche MJ, van der Wouden JC, Uijen JH, *et al.* Randomised placebo-controlled trial of inhaled sodium cromoglycate in 1–4-year-old children with moderate asthma. *Lancet* 1997; 350: 1060–1064.
26. O'Byrne PM, Barnes PJ, Rodriguez-Roisin R, *et al.* Low dose inhaled budesonide and formoterol in mild persistent asthma: the OPTIMA randomized trial. *Am J Respir Crit Care Med* 2001; 164: 1392–1397.
27. Sorkness CA, Lemanske RF Jr, Mauger DT, *et al.* Long-term comparison of 3 controller regimens for mild-moderate persistent childhood asthma: the Pediatric Asthma Controller Trial. *J Allergy Clin Immunol* 2007; 119: 64–72.
28. Covar RA, Szefler SJ, Zeiger RS, *et al.* Factors associated with asthma exacerbations during a long-term clinical trial of controller medications in children. *J Allergy Clin Immunol* 2008; 122: 741–747.
29. Morgan WJ, Crain EF, Gruchalla RS, *et al.* Results of a home-based environmental intervention among urban children with asthma. *N Engl J Med* 2004; 351: 1068–1080.
30. Juniper EF, O'Byrne PM, Guyatt GH, *et al.* Development and validation of a questionnaire to measure asthma control. *Eur Respir J* 1999; 14: 902–907.
31. Nathan RA, Sorkness CA, Kosinski M, *et al.* Development of the asthma control test: a survey for assessing asthma control. *J Allergy Clin Immunol* 2004; 113: 59–65.
32. O'Byrne PM, Reddel HK, Eriksson G, *et al.* Measuring asthma control: a comparison of three classification systems. *Eur Respir J* 2010; 36: 269–276.
33. Juniper EF, Bousquet J, Abetz L, *et al.* Identifying 'well-controlled' and 'not well-controlled' asthma using the Asthma Control Questionnaire. *Respir Med* 2006; 100: 616–621.
34. Juniper EF, Svensson K, Mörk AC, *et al.* Measurement properties and interpretation of three shortened versions of the asthma control questionnaire. *Respir Med* 2005; 99: 553–558.
35. Schatz M, Kosinski M, Yarlas AS, *et al.* The minimally important difference of the Asthma Control Test. *J Allergy Clin Immunol* 2009; 124: 719–723.
36. Schatz M, Sorkness CA, Li JT, *et al.* Asthma Control Test: reliability, validity, and responsiveness in patients not previously followed by asthma specialists. *J Allergy Clin Immunol* 2006; 117: 549–556.
37. Jia CE, Zhang HP, Lv Y, *et al.* The Asthma Control Test and Asthma Control Questionnaire for assessing asthma control: systematic review and meta-analysis. *J Allergy Clin Immunol* 2013; 131: 695–703.
38. Wildfire JJ, Gergen PJ, Sorkness CA, *et al.* Development and validation of the Composite Asthma Severity Index – an outcome measure for use in children and adolescents. *J Allergy Clin Immunol* 2012; 129: 694–701.
39. O'Byrne PM, Bisgaard H, Godard PP, *et al.* Budesonide/formoterol combination therapy as both maintenance and reliever medication in asthma. *Am J Respir Crit Care Med* 2005; 171: 129–136.

40. Rand CS, Wright RJ, Cabana MD, *et al.* Mediators of asthma outcomes. *J Allergy Clin Immunol* 2012; 129: Suppl. 3, S136–S141.
41. Erickson SR, Kirking DM. Variation in the distribution of patient-reported outcomes based on different definitions of defining asthma severity. *Curr Med Res Opin* 2004; 20: 1863–1872.
42. Juniper EF, Guyatt GH, Epstein RS, *et al.* Evaluation of impairment of health related quality of life in asthma: development of a questionnaire for use in clinical trials. *Thorax* 1992; 47: 76–83.
43. Juniper EF, Buist AS, Cox FM, *et al.* Validation of a standardized version of the Asthma Quality of Life Questionnaire. *Chest* 1999; 115: 1265–1270.
44. Sanjuás C, Alonso J, Prieto L, *et al.* Health-related quality of life in asthma: a comparison between the St George's Respiratory Questionnaire and the Asthma Quality of Life Questionnaire. *Qual Life Res* 2002; 11: 729–738.
45. Juniper EF, Guyatt GH, Cox FM, *et al.* Development and validation of the Mini Asthma Quality of Life Questionnaire. *Eur Respir J* 1999; 14: 32–38.
46. Marks GB, Dunn SM, Woolcock AJ. A scale for the measurement of quality of life in adults with asthma. *J Clin Epidemiol* 1992; 45: 461–472.
47. Katz PP, Eisner MD, Henke J, *et al.* The Marks Asthma Quality of Life Questionnaire: further validation and examination of responsiveness to change. *J Clin Epidemiol* 1999; 52: 667–675.
48. Jones PW, Quirk FH, Baveystock CM, *et al.* A self-complete measure of health status for chronic airflow limitation. The St. George's Respiratory Questionnaire. *Am Rev Respir Dis* 1992; 145: 1321–1327.
49. Barley EA, Quirk FH, Jones PW. Asthma health status measurement in clinical practice: validity of a new short and simple instrument. *Respir Med* 1998; 92: 1207–1214.
50. Singh AM, Busse WW. Asthma exacerbations. 2: aetiology. *Thorax* 2006; 61: 809–816.
51. Busse WW, Lemanske RF Jr, Gern JE. Role of viral respiratory infections in asthma and asthma exacerbations. *Lancet* 2010; 376: 826–834.
52. Rank MA, Hagan JB, Park MA, *et al.* The risk of asthma exacerbation after stopping low-dose inhaled corticosteroids: a systematic review and meta-analysis of randomized controlled trials. *J Allergy Clin Immunol* 2013; 131: 724–729.
53. Fuhlbrigge A, Peden D, Apter AJ, *et al.* Asthma outcomes: exacerbations. *J Allergy Clin Immunol* 2012; 129: Suppl. 3, S34–S48.
54. O'Byrne PM. Therapeutic strategies to reduce asthma exacerbations. *J Allergy Clin Immunol* 2011; 128: 257–263.
55. Pauwels RA, Löfdahl CG, Postma DS, *et al.* Effect of inhaled formoterol and budesonide on exacerbations of asthma. Formoterol and Corticosteroids Establishing Therapy (FACET) International Study Group. *N Engl J Med* 1997; 337: 1405–1411.
56. Humbert M, Beasley R, Ayres J, *et al.* Benefits of omalizumab as add-on therapy in patients with severe persistent asthma who are inadequately controlled despite best available therapy (GINA 2002 step 4 treatment): INNOVATE. *Allergy* 2005; 60: 309–316.
57. Ayres JG, Higgins B, Chilvers ER, *et al.* Efficacy and tolerability of anti-immunoglobulin E therapy with omalizumab in patients with poorly controlled (moderate-to-severe) allergic asthma. *Allergy* 2004; 59: 701–708.
58. Busse W, Corren J, Lanier BQ, *et al.* Omalizumab, anti-IgE recombinant humanized monoclonal antibody, for the treatment of severe allergic asthma. *J Allergy Clin Immunol* 2001; 108: 184–190.
59. Solèr M, Matz J, Townley R, *et al.* The anti-IgE antibody omalizumab reduces exacerbations and steroid requirement in allergic asthmatics. *Eur Respir J* 2001; 18: 254–261.
60. Hanania NA, Alpan O, Hamilos DL, *et al.* Omalizumab in severe allergic asthma inadequately controlled with standard therapy: a randomized trial. *Ann Intern Med* 2011; 154: 573–582.
61. Haldar P, Brightling CE, Hargadon B, *et al.* Mepolizumab and exacerbations of refractory eosinophilic asthma. *N Engl J Med* 2009; 360: 973–984.
62. Nair P, Pizzichini MM, Kjarsgaard M, *et al.* Mepolizumab for prednisone-dependent asthma with sputum eosinophilia. *N Engl J Med* 2009; 360: 985–993.
63. Pavord ID, Korn S, Howarth P, *et al.* Mepolizumab for severe eosinophilic asthma (DREAM): a multicentre, double-blind, placebo-controlled trial. *Lancet* 2012; 380: 651–659.
64. Corren J, Lemanske RF, Hanania NA, *et al.* Lebrikizumab treatment in adults with asthma. *N Engl J Med* 2011; 365: 1088–1098.
65. Wenzel S, Ford L, Pearlman D, *et al.* Dupilumab in persistent asthma with elevated eosinophil levels. *N Engl J Med* 2013; 368: 2455–2466.
66. Sears MR, Radner F. Safety of formoterol in asthma clinical trials: an update. *Eur Respir J* 2013 [In press DOI: 10.1183/09031936.00004713].
67. Nelson HS, Weiss ST, Bleecker ER, *et al.* The Salmeterol Multicenter Asthma Research Trial: a comparison of usual pharmacotherapy for asthma or usual pharmacotherapy plus salmeterol. *Chest* 2006; 129: 15–26.
68. Castle W, Fuller R, Hall J, *et al.* Serevent nationwide surveillance study: comparison of salmeterol with salbutamol in asthmatic patients who require regular bronchodilator treatment. *BMJ* 1993; 306: 1034–1037.
69. Suissa S, Ariel A. US Food and Drug Administration-mandated trials of long-acting β-agonists safety in asthma: will we know the answer? *Chest* 2013; 143: 1208–1213.

70. Chowdhury BA, Seymour SM, Levenson MS. Assessing the safety of adding LABAs to inhaled corticosteroids for treating asthma. *N Engl J Med* 2011; 364: 2473–2475.
71. Travers AH, Jones AP, Camargo CA Jr, *et al.* Intravenous β_2-agonists *versus* intravenous aminophylline for acute asthma. *Cochrane Database Syst Rev* 2012; 12: CD010256.
72. Travers AH, Milan SJ, Jones AP, *et al.* Addition of intravenous β_2-agonists to inhaled β_2-agonists for acute asthma. *Cochrane Database Syst Rev* 2012; 12: CD010179.
73. Lim WJ, Mohammed AR, Carson KV, *et al.* Non-invasive positive pressure ventilation for treatment of respiratory failure due to severe acute exacerbations of asthma. *Cochrane Database Syst Rev* 2012; 12: CD004360.
74. Powell C, Dwan K, Milan SJ, *et al.* Inhaled magnesium sulfate in the treatment of acute asthma. *Cochrane Database Syst Rev* 2012; 12: CD003898.
75. Jat KR, Chawla D. Ketamine for management of acute exacerbations of asthma in children. *Cochrane Database Syst Rev* 2012; 11: CD009293.
76. Miller MR, Hankinson J, Brusasco V, *et al.* Standardisation of spirometry. *Eur Respir J* 2005; 26: 319–338.
77. Tepper RS, Wise RS, Covar R, *et al.* Asthma outcomes: pulmonary physiology. *J Allergy Clin Immunol* 2012; 129: Suppl. 3, S65–S87.
78. Quanjer PH, Tammeling GJ, Cotes JE, *et al.* Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. *Eur Respir J* 1993; 6: Suppl. 16, 5–40.
79. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med* 1999; 159: 179–187.
80. Quanjer PH, Stanojevic S, Cole TJ, *et al.* Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012; 40: 1324–1343.
81. Tashkin DP, Celli B, Decramer M, *et al.* Bronchodilator responsiveness in patients with COPD. *Eur Respir J* 2008; 31: 742–750.
82. James AL, Wenzel S. Clinical relevance of airway remodelling in airway diseases. *Eur Respir J* 2007; 30: 134–155.
83. Frey U, Suki B. Complexity of chronic asthma and chronic obstructive pulmonary disease: implications for risk assessment, and disease progression and control. *Lancet* 2008; 372: 1088–1099.
84. Frey U, Brodbeck T, Majumdar A, *et al.* Risk of severe asthma episodes predicted from fluctuation analysis of airway function. *Nature* 2005; 438: 667–670.
85. Côté J, Cartier A, Malo JL, *et al.* Compliance with peak expiratory flow monitoring in home management of asthma. *Chest* 1998; 113: 968–972.
86. Reddel HK, Toelle BG, Marks GB, *et al.* Analysis of adherence to peak flow monitoring when recording of data is electronic. *BMJ* 2002; 324: 146–147.
87. Lommatzsch M. Airway hyperresponsiveness: new insights into the pathogenesis. *Semin Respir Crit Care Med* 2012; 33: 579–587.
88. Cockcroft DW. Direct challenge tests: airway hyperresponsiveness in asthma: its measurement and clinical significance. *Chest* 2010; 138: Suppl. 2, 18S–24S.
89. Anderson SD. Indirect challenge tests: airway hyperresponsiveness in asthma: its measurement and clinical significance. *Chest* 2010; 138: Suppl. 2, 25S–30S.
90. Koskela HO, Hyvärinen L, Brannan JD, *et al.* Responsiveness to three bronchial provocation tests in patients with asthma. *Chest* 2003; 124: 2171–2177.
91. Koskela HO, Hyvärinen L, Brannan JD, *et al.* Sensitivity and validity of three bronchial provocation tests to demonstrate the effect of inhaled corticosteroids in asthma. *Chest* 2003; 124: 1341–1349.
92. Van Den Berge M, Meijer RJ, Kerstjens HA, *et al.* PC_{20} adenosine 5′-monophosphate is more closely associated with airway inflammation in asthma than PC_{20} methacholine. *Am J Respir Crit Care Med* 2001; 163: 1546–1550.
93. Suh DI, Lee JK, Kim CK, *et al.* Bronchial hyperresponsiveness to methacholine and adenosine 5′-monophosphate, and the presence and degree of atopy in young children with asthma. *Clin Exp Allergy* 2011; 41: 338–345.
94. Crapo RO, Casaburi R, Coates AL, *et al.* Guidelines for methacholine and exercise challenge testing – 1999. This official statement of the American Thoracic Society was adopted by the ATS Board of Directors, July 1999. *Am J Respir Crit Care Med* 2000; 161: 309–329.
95. Schulze J, Rosewich M, Riemer C, *et al.* Methacholine challenge – comparison of an ATS protocol to a new rapid single concentration technique. *Respir Med* 2009; 103: 1898–1903.
96. Merget R, Jörres RA, Heinze E, *et al.* Development of a 1-concentration-4-step dosimeter protocol for methacholine testing. *Respir Med* 2009; 103: 607–613.
97. Sont JK, Willems LN, Bel EH, *et al.* Clinical control and histopathologic outcome of asthma when using airway hyperresponsiveness as an additional guide to long-term treatment. The AMPUL Study Group. *Am J Respir Crit Care Med* 1999; 159: 1043–1051.
98. Grazzini M, Scano G, Foglio K, *et al.* Relevance of dyspnoea and respiratory function measurements in monitoring of asthma: a factor analysis. *Respir Med* 2001; 95: 246–250.
99. Stanbrook MB, Chapman KR, Kesten S. Gas trapping as a predictor of positive methacholine challenge in patients with normal spirometry results. *Chest* 1995; 107: 992–995.
100. Magnussen H, Richter K, Taube C. Are chronic obstructive pulmonary disease (COPD) and asthma different diseases? *Clin Exp Allergy* 1998; 28: Suppl. 5, 187–194.

101. Deesomchok A, Fisher T, Webb KA, *et al.* Effects of obesity on perceptual and mechanical responses to bronchoconstriction in asthma. *Am J Respir Crit Care Med* 2010; 181: 125–133.
102. Sorkness RL, Bleecker ER, Busse WW, *et al.* Lung function in adults with stable but severe asthma: air trapping and incomplete reversal of obstruction with bronchodilation. *J Appl Physiol* 2008; 104: 394–403.
103. Criée CP, Sorichter S, Smith HJ, *et al.* Body plethysmography – its principles and clinical use. *Respir Med* 2011; 105: 959–971.
104. Tantucci C, Guerini M, Boni E, *et al.* Tidal airway closure during bronchoconstriction in asthma: usefulness of lung volume measurements. *J Asthma* 2011; 48: 33–40.
105. Kaminsky DA, Wenzel SE, Carcano C, *et al.* Hyperpnea-induced changes in parenchymal lung mechanics in normal subjects and in asthmatics. *Am J Respir Crit Care Med* 1997; 155: 1260–1266.
106. Oostveen E, MacLeod D, Lorino H, *et al.* The forced oscillation technique in clinical practice: methodology, recommendations and future developments. *Eur Respir J* 2003; 22: 1026–1041.
107. Koopmans JG, Lutter R, Jansen HM, *et al.* Clinically masked increases in bronchial inflammation in guideline-treated persistent asthma. *Pulm Pharmacol Ther* 2006; 19: 397–403.
108. Dompeling E, van Schayck CP, van Grunsven PM, *et al.* Slowing the deterioration of asthma and chronic obstructive pulmonary disease observed during bronchodilator therapy by adding inhaled corticosteroids. A 4-year prospective study. *Ann Intern Med* 1993; 118: 770–778.
109. Reed CE. The natural history of asthma in adults: the problem of irreversibility. *J Allergy Clin Immunol* 1999; 103: 539–547.
110. Selroos O, Pietinalho A, Löfroos AB, *et al.* Effect of early *vs* late intervention with inhaled corticosteroids in asthma. *Chest* 1995; 108: 1228–1234.
111. Wenzel SE. Asthma phenotypes: the evolution from clinical to molecular approaches. *Nat Med* 2012; 18: 716–725.
112. McGrath KW, Icitovic N, Boushey HA, *et al.* A large subgroup of mild-to-moderate asthma is persistently noneosinophilic. *Am J Respir Crit Care Med* 2012; 185: 612–619.
113. Moore WC, Meyers DA, Wenzel SE, *et al.* Identification of asthma phenotypes using cluster analysis in the Severe Asthma Research Program. *Am J Respir Crit Care Med* 2010; 181: 315–323.
114. Haldar P, Pavord ID, Shaw DE, *et al.* Cluster analysis and clinical asthma phenotypes. *Am J Respir Crit Care Med* 2008; 178: 218–224.
115. Woodruff PG, Modrek B, Choy DF, *et al.* T-helper type 2-driven inflammation defines major subphenotypes of asthma. *Am J Respir Crit Care Med* 2009; 180: 388–395.
116. Szefler SJ, Wenzel S, Brown R, *et al.* Asthma outcomes: biomarkers. *J Allergy Clin Immunol* 2012; 129: Suppl. 3, S9–S23.
117. ATS/ERS recommendations for standardized procedures for the online and offline measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide, 2005. *Am J Respir Crit Care Med* 2005; 171: 912–930.
118. Dweik RA, Boggs PB, Erzurum SC, *et al.* An official ATS clinical practice guideline: interpretation of exhaled nitric oxide levels (FeNO) for clinical applications. *Am J Respir Crit Care Med* 2011; 184: 602–615.
119. Hastie AT, Moore WC, Li H, *et al.* Biomarker surrogates do not accurately predict sputum eosinophil and neutrophil percentages in asthmatic subjects. *J Allergy Clin Immunol* 2013; 130: 510–515.
120. Cowan DC, Cowan JO, Palmay R, *et al.* Effects of steroid therapy on inflammatory cell subtypes in asthma. *Thorax* 2010; 65: 384–390.
121. Boushey HA, Sorkness CA, King TS, *et al.* Daily *versus* as-needed corticosteroids for mild persistent asthma. *N Engl J Med* 2005; 352: 1519–1528.
122. Claman DM, Boushey HA, Liu J, *et al.* Analysis of induced sputum to examine the effects of prednisone on airway inflammation in asthmatic subjects. *J Allergy Clin Immunol* 1994; 94: 861–869.
123. van Rensen EL, Straathof KC, Veselic-Charvat MA, *et al.* Effect of inhaled steroids on airway hyperresponsiveness, sputum eosinophils, and exhaled nitric oxide levels in patients with asthma. *Thorax* 1999; 54: 403–408.
124. Djukanović R, Wilson SJ, Kraft M, *et al.* Effects of treatment with anti-immunoglobulin E antibody omalizumab on airway inflammation in allergic asthma. *Am J Respir Crit Care Med* 2004; 170: 583–593.
125. Holgate ST, Chuchalin AG, Hébert J, *et al.* Efficacy and safety of a recombinant anti-immunoglobulin E antibody (omalizumab) in severe allergic asthma. *Clin Exp Allergy* 2004; 34: 632–638.
126. Massanari M, Holgate ST, Busse WW, *et al.* Effect of omalizumab on peripheral blood eosinophilia in allergic asthma. *Respir Med* 2010; 104: 188–196.
127. Leckie MJ, ten Brinke A, Khan J, *et al.* Effects of an interleukin-5 blocking monoclonal antibody on eosinophils, airway hyper-responsiveness, and the late asthmatic response. *Lancet* 2000; 356: 2144–2148.
128. Liu Y, Zhang S, Li DW, *et al.* Efficacy of anti-interleukin-5 therapy with mepolizumab in patients with asthma: a meta-analysis of randomized placebo-controlled trials. *PLoS One* 2013; 8: e59872.
129. Piper E, Brightling C, Niven R, *et al.* A phase II placebo-controlled study of tralokinumab in moderate-to-severe asthma. *Eur Respir J* 2013; 41: 330–338.
130. Karrasch S, Ernst K, Behr J, *et al.* Exhaled nitric oxide and influencing factors in a random population sample. *Respir Med* 2011; 105: 713–718.
131. Berry MA, Shaw DE, Green RH, *et al.* The use of exhaled nitric oxide concentration to identify eosinophilic airway inflammation: an observational study in adults with asthma. *Clin Exp Allergy* 2005; 35: 1175–1179.

132. Knuffman JE, Sorkness CA, Lemanske RF Jr, *et al.* Phenotypic predictors of long-term response to inhaled corticosteroid and leukotriene modifier therapies in pediatric asthma. *J Allergy Clin Immunol* 2009; 123: 411–416.
133. Smith AD, Cowan JO, Brassett KP, *et al.* Exhaled nitric oxide: a predictor of steroid response. *Am J Respir Crit Care Med* 2005; 172: 453–459.
134. Szefler SJ, Martin RJ, King TS, *et al.* Significant variability in response to inhaled corticosteroids for persistent asthma. *J Allergy Clin Immunol* 2002; 109: 410–418.
135. Hanania NA, Wenzel S, Rosén K, *et al.* Exploring the effects of omalizumab in allergic asthma: an analysis of biomarkers in the EXTRA study. *Am J Respir Crit Care Med* 2013; 187: 804–811.
136. Smith AD, Cowan JO, Brassett KP, *et al.* Use of exhaled nitric oxide measurements to guide treatment in chronic asthma. *N Engl J Med* 2005; 352: 2163–2173.
137. Calhoun WJ, Ameredes BT, King TS, *et al.* Comparison of physician-, biomarker-, and symptom-based strategies for adjustment of inhaled corticosteroid therapy in adults with asthma: the BASALT randomized controlled trial. *JAMA* 2012; 308: 987–997.
138. Powell H, Murphy VE, Taylor DR, *et al.* Management of asthma in pregnancy guided by measurement of fraction of exhaled nitric oxide: a double-blind, randomised controlled trial. *Lancet* 2011; 378: 983–990.
139. Paggiaro PL, Chanez P, Holz O, *et al.* Sputum induction. *Eur Respir J* 2002; 20: Suppl. 37, 3S–8S.
140. Haldar P, Pavord ID. Noneosinophilic asthma: a distinct clinical and pathologic phenotype. *J Allergy Clin Immunol* 2007; 119: 1043–1052.
141. Petsky HL, Cates CJ, Li AM, *et al.* Tailored interventions based on exhaled nitric oxide *versus* clinical symptoms for asthma in children and adults. *Cochrane Database Syst Rev* 2008; CD006340.
142. Petsky HL, Kynaston JA, Turner C, *et al.* Tailored interventions based on sputum eosinophils *versus* clinical symptoms for asthma in children and adults. *Cochrane Database Syst Rev* 2007; CD005603.
143. Petsky HL, Cates CJ, Lasserson TJ, *et al.* A systematic review and meta-analysis: tailoring asthma treatment on eosinophilic markers (exhaled nitric oxide or sputum eosinophils). *Thorax* 2012; 67: 199–208.
144. Burney P, Malmberg E, Chinn S, *et al.* The distribution of total and specific serum IgE in the European Community Respiratory Health Survey. *J Allergy Clin Immunol* 1997; 99: 314–322.
145. Worldwide variation in prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and atopic eczema: ISAAC. The International Study of Asthma and Allergies in Childhood (ISAAC) Steering Committee. *Lancet* 1998; 351: 1225–1232.
146. Bousquet J, Wenzel S, Holgate S, *et al.* Predicting response to omalizumab, an anti-IgE antibody, in patients with allergic asthma. *Chest* 2004; 125: 1378–1386.
147. Bousquet J, Rabe K, Humbert M, *et al.* Predicting and evaluating response to omalizumab in patients with severe allergic asthma. *Respir Med* 2007; 101: 1483–1492.
148. Wahn U, Martin C, Freeman P, *et al.* Relationship between pretreatment specific IgE and the response to omalizumab therapy. *Allergy* 2009; 64: 1780–1787.
149. Woodruff PG, Boushey HA, Dolganov GM, *et al.* Genome-wide profiling identifies epithelial cell genes associated with asthma and with treatment response to corticosteroids. *Proc Natl Acad Sci USA* 2007; 104: 15858–15863.
150. Jia G, Erickson RW, Choy DF, *et al.* Periostin is a systemic biomarker of eosinophilic airway inflammation in asthmatic patients. *J Allergy Clin Immunol* 2012; 130: 647–654.
151. Irwin RS. The role of conflict of interest in reporting of scientific information. *Chest* 2009; 136: 253–259.
152. Sismondo S. Pharmaceutical company funding and its consequences: a qualitative systematic review. *Contemp Clin Trials* 2008; 29: 109–113.
153. Turner EH, Matthews AM, Linardatos E, *et al.* Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* 2008; 358: 252–260.
154. Als-Nielsen B, Chen W, Gluud C, *et al.* Association of funding and conclusions in randomized drug trials: a reflection of treatment effect or adverse events? *JAMA* 2003; 290: 921–928.
155. Bond K, Spooner C, Tjosvold L, *et al.* The nature and influence of pharmaceutical industry involvement in asthma trials. *Can Respir J* 2012; 19: 267–271.
156. Sun X, Briel M, Busse JW, *et al.* The influence of study characteristics on reporting of subgroup analyses in randomised controlled trials: systematic review. *BMJ* 2011; 342: d1569.
157. De AC, Drazen JM, Frizelle FA, *et al.* Clinical trial registration: a statement from the International Committee of Medical Journal Eds. *Ann Intern Med* 2004; 141: 477–478.
158. Laine C, Horton R, DeAngelis CD, *et al.* Clinical trial registration – looking back and moving ahead. *N Engl J Med* 2007; 356: 2734–2736.
159. Zarin DA, Keselman A. Registering a clinical trial in ClinicalTrials.gov. *Chest* 2007; 131: 909–912.
160. Drazen JM, de Leeuw PW, Laine C, *et al.* Toward more uniform conflict disclosures – the updated ICMJE conflict of interest reporting form. *N Engl J Med* 2010; 363: 188–189.
161. Herland K, Akselsen JP, Skjønsberg OH, *et al.* How representative are clinical study patients with asthma or COPD for a larger "real life" population of patients with obstructive lung disease? *Respir Med* 2005; 99: 11–19.
162. Price D, Hillyer EV, van der Molen T. Efficacy *versus* effectiveness trials: informing guidelines for asthma management. *Curr Opin Allergy Clin Immunol* 2013; 13: 50–57.

163. Colice G, Martin RJ, Israel E, *et al.* Asthma outcomes and costs of therapy with extrafine beclomethasone and fluticasone. *J Allergy Clin Immunol* 2013; 132: 45–54.
164. Price D, Musgrave SD, Shepstone L, *et al.* Leukotriene antagonists as first-line or add-on asthma-controller therapy. *N Engl J Med* 2011; 364: 1695–1707.
165. Chauhan BF, Ducharme FM. Anti-leukotriene agents compared to inhaled corticosteroids in the management of recurrent and/or chronic asthma in adults and children. *Cochrane Database Syst Rev* 2012; 5: CD002314.
166. Dahlén SE, Dahlén B, Drazen JM. Asthma treatment guidelines meet the real world. *N Engl J Med* 2011; 364: 1769–1770.

Chapter 3

Idiopathic pulmonary fibrosis

Toby M. Maher

SUMMARY: Following publication of the first true multicentre randomised controlled trial in 2004, there has been an exponential growth in clinical trial activity in idiopathic pulmonary fibrosis (IPF). Whilst the majority of trials published to date have been negative, each has generated information on the strengths and limitations of a range of clinical trial outcome measures. Furthermore, analysis of clinical trial data has improved understanding of disease behaviour as well as the utility of many routinely collected clinical measures. However, despite this, there remains considerable debate about the optimal choice of primary end-point for registration studies in IPF with all-cause mortality, progression-free survival and forced vital capacity all having their advocates. For early-phase studies in IPF there is a lot less clarity around optimal outcome measures with few if any able to detect meaningful disease change in a sufficiently small sample size.

NIHR Respiratory Biomedical Research Unit, Royal Brompton Hospital, London; Centre for Respiratory Research, University College London, Rayne Institute, London; and National Heart and Lung Institute, Imperial College London, London, UK.

Correspondence: T.M. Maher, Interstitial Lung Disease Unit, Royal Brompton Hospital, Sydney Street, London, SW3 6NP, UK.
Email: t.maher@rbht.nhs.uk

Eur Respir Monogr 2013; 62: 37–53.
Copyright ERS 2013.
DOI: 10.1183/1025448x.10044512
Print ISBN: 978-1-84984-044-6
Online ISBN: 978-1-84984-045-3
Print ISSN: 1025-448x
Online ISSN: 2075-6674

Frequently considered by pulmonologists to be untreatable, idiopathic pulmonary fibrosis (IPF) has lagged behind other respiratory diseases when it comes to the development and delivery of clinical trials. After a belated infancy, beginning in the late 1990s, IPF clinical trial design, at least when it comes to late-phase studies, is finally reaching maturity. Notwithstanding recent heated debates, concerning optimal primary end-points for registration studies [1, 2], the accumulated body of knowledge gained through the delivery of multiple late-phase clinical trials has led to a greatly improved understanding of the value, and limitations, of various outcome measures that have been applied as study end-points. The last decade has seen a massive increase in clinical trial activity with the number of patients being randomised into multicentre, controlled trails increasing almost twentyfold compared to the decade before [3]. The culmination of these studies, thus far, has been the licensing in Europe, Japan and Asia of the novel, first-in-class IPF therapy, pirfenidone [4]. Despite this success there remain many challenges and controversies in the field of IPF clinical trials. Whilst some of these are common to other disease and disciplines, a number are unique to IPF. These issues, as they relate to the measurement of outcomes, will form the focus of discussion in this manuscript.

Epidemiology and pathogenesis

IPF is an inexorably progressive and invariably fatal disease characterised by the accumulation of collagen rich extracellular matrix within the alveolar interstitium, which ultimately results in

architectural destruction of the lung with replacement of alveoli by mucin-filled honeycomb cysts [5]. Median survival, albeit in data derived before the emergence of effective therapies, is 2.8–4.2 years from diagnosis [6–8]. Based on studies from both the UK and the USA the condition appears to be increasing in incidence [9, 10]. In the UK IPF is responsible for the deaths of 5000 people per year. IPF is more common in males whilst 80% of sufferers have a past smoking history. Various case-control studies suggest that past exposure to work-place dusts, be they organic, metal or mineral, results in an increased odds ratio for developing IPF. Similarly, a past history of specific viral infection and gastro-oesophageal reflux has been linked to the development of IPF [11, 12].

It has long been recognised that somewhere between 1% and 5% of IPF cases are familial, that is to say they occur in individuals with a first-degree relative who also has the disease [13]. Over the last decade a number of key genetic studies, undertaken in sporadic and familial cases of IPF, have identified a number of potentially causative gene polymorphisms. Single-nucleotide polymorphisms have been found in the gene for surfactant protein C [14], the gene for the mucin MUC5B [15] and in genes for enzymes important in the maintenance of telomere length [16]. Two recently published, large-scale genome-wide association studies have confirmed these previously identified polymorphisms as risk factors for the development of IPF and have also identified other genes conferring susceptibility to IPF [17, 18]. These genes appear to be predominantly involved in regulating host defence and ensuring epithelial integrity.

These genetic and epidemiological observations, combined with experimental preclinical studies strongly suggest that IPF arises as the consequence of an aberrant wound healing response driven by recurrent epithelial injury in genetically susceptible individuals (fig. 1) [5]. The normal wound healing response involves the interaction of a multitude of carefully orchestrated and overlapping pathways that act to restore epithelial integrity and, where possible, normal tissue structure and function following injury. In the lungs of individuals with IPF all of these same pathways are activated, but for reasons, which remain to be fully elucidated, their activation fails to restore epithelial barrier integrity. Instead, the continued synthesis and activation of a host of profibrotic mediators ultimately results in progressive expansion of the extracellular matrix and destruction of normal lung architecture. Although incomplete, improved understanding of the pathogenesis of IPF has identified a large number of specific therapeutic targets [19]. This has led, in the last few years, to a rapid expansion in the number of compounds entering early-phase clinical trials in IPF.

The diagnosis and treatment of IPF

The diagnosis of IPF, formalised in international guidelines first published in 2002 [20] and recently revised in 2011 [21], rests on radiological or histological findings of usual interstitial pneumonia (UIP) in the absence of an identifiable alternative cause. Approximately two thirds of patients with UIP/IPF have a characteristic appearance on high-resolution computed tomography (HRCT), consisting subpleural reticular change and honeycombing in a predominantly basal distribution with a paucity of ground-glass attenuation. For this reason biopsy is only required to confirm a diagnosis of IPF in the minority of individuals. Because, a diagnosis of IPF rests on the integration of clinical and radiographic data, together with, in a proportion of cases, histology, current guidelines recommend that a final diagnosis is only assigned following multidisciplinary discussion involving a radiologist, pathologist and specialist interstitial lung disease (ILD) clinician [22]. Given the importance of these three facets of the diagnostic process, the majority of late-phase IPF clinical trials have now moved to centralised assessment of computed tomography (CT) scans and histology as part of the screening process for potential study candidates.

Publication of the first American Thoracic Society (ATS)/European Respiratory Society (ERS) guidelines was an important prerequisite to the development of clinical trials for IPF [20]. Prior to this, there were major international variations in the nomenclature and criteria used to assign a diagnosis in individuals with fibrotic ILD. Even in the USA, where IPF was the preferred diagnostic term, before 2002 the diagnostic label of IPF was used far more loosely to refer to a

range of ILDs, many of which have subsequently been shown to differ in their underlying pathobiology. Harmonisation of diagnostic criteria ensured that studies of IPF contained a relatively homogeneous population of subjects with, as far as possible, shared disease behaviour. This in turn made it possible for the development of international multicentre studies. Before 2002, published clinical trials in IPF were few and far between and tended to be small in size and frequently single-centre. It has been calculated that globally only 114 individuals with IPF were enrolled in clinical trials during the 1990s (fig. 2) [37]. The first true multicentre, international IPF trial was published as recently as 2004 [28] and this heralded a near twenty-fold increase in IPF patients enrolled in trials over the remainder of that decade. Since the turn of the current decade this increase in IPF trial activity has been sustained and promises to continue rising. This growth has been driven both by increased pharmaceutical company involvement in the field but also, importantly, the development of clinical trial networks in both the USA and Europe. Whilst the majority of studies published to date have been negative, all have contributed important knowledge, regarding optimal trial design and outcome measurement, for IPF clinical studies.

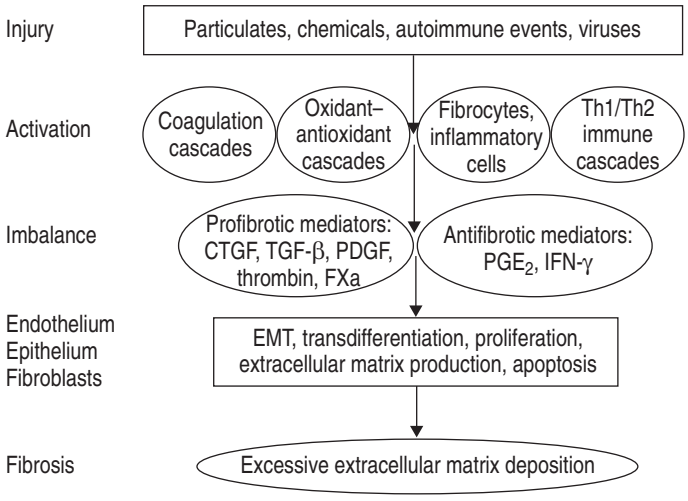


Figure 1. Current understanding of the pathogenesis of idiopathic pulmonary fibrosis (IPF) suggests that repetitive alveolar epithelial injury results in basement membrane denudation and activation of key pathways involved in the wound healing response. This in turn leads to fibroblast proliferation, transformation of fibroblasts to myofibroblasts and expansion of the extracellular matrix. These effects are augmented by the influx of circulating inflammatory cells, including the putative bone marrow-derived fibroblast precursor, the fibrocyte. Various treatments are in development targeting different aspects of IPF disease pathogenesis, through inhibition of fibrogenesis, promotion of antifibrotic pathways, or through the reduction of alveolar injury. Th: T-helper cell; CTGF: connective tissue growth factor; TGF: transforming growth factor; PDGF: platelet-derived growth factor; FXa: factor Xa; PGE_2: prostaglandin E_2; IFN: interferon; EMT: epithelial–mesenchymal transition. Reproduced from [5].

Challenges in IPF trial design

Before reviewing the ins and outs of the various possible IPF clinical trial outcome measures that have been used, it is worth considering the challenges inherent in identifying and assessing the efficacy of potential therapeutic compounds for IPF.

Preclinical target and compound validation

The molecular mechanisms driving the pathogenesis of IPF are complex and the many pathways involved demonstrate considerable redundancy and pleiotrophism. Advances in the understanding of the basic mechanisms driving the development of fibrosis have identified a plethora of potential therapeutic targets. Whilst this presents a major opportunity for drug development it also brings several challenges. The identification of appropriate preclinical models of disease has lagged a long way behind the evolution of late-phase clinical trials. The rodent bleomycin model of pulmonary fibrosis has been widely used to validate candidate drugs for IPF. Yet, as has been described elsewhere, the early phase of the bleomycin model is one of an inflammatory acute lung injury with widespread epithelial apoptosis [38]. As a consequence, therapies given before, or when bleomycin is administered will, if anti-inflammatory or anti-apoptotic, give the false appearance of

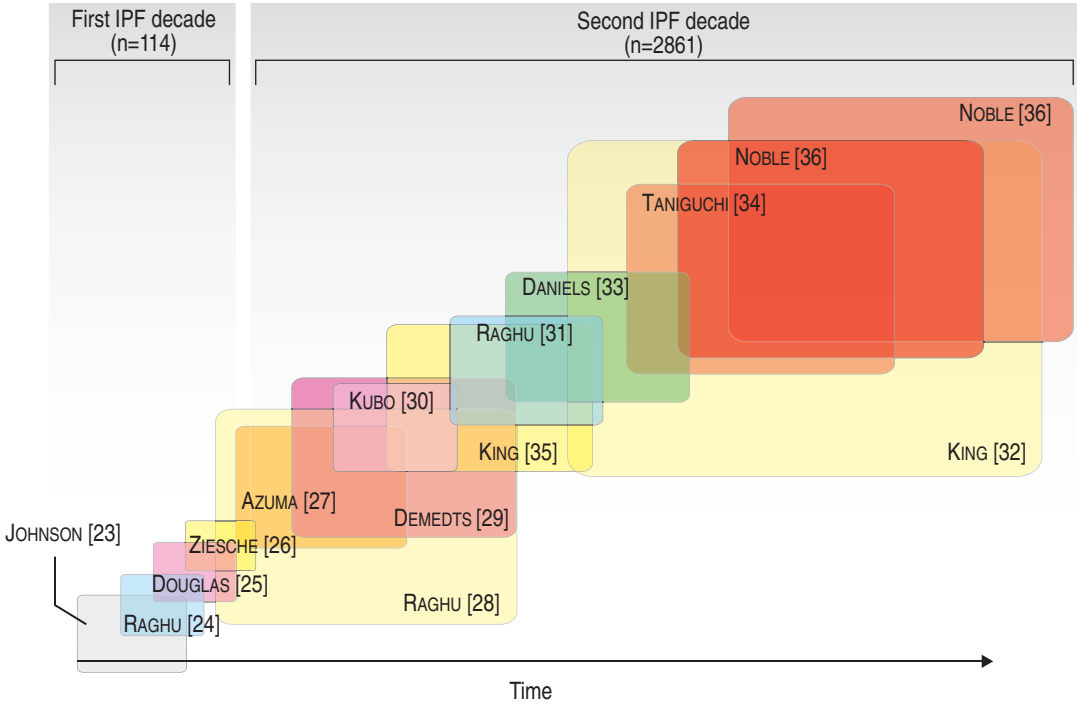


Figure 2. The number of patients enrolled in clinical studies for idiopathic pulmonary fibrosis (IPF), has increased substantially in the past decade when compared with the previous decade. Reproduced from [3].

being antifibrotic. Many of the therapies tested in early IPF clinical trials were done so, in part, on the basis of inaccurate preclinical validation. Imatinib, the tyrosine kinase inhibitor, provides just such an example [33]. Early published studies suggested that imatinib had antifibrotic effects *in vitro* and *in vivo* in mice [39]. Yet closer inspection of the animal studies reveals that imatinib only blocked the development of fibrosis in mice when administered prophylactically (*i.e.* at the same time as the bleomycin). When imatinib was given to mice 7 days post-bleomycin (a point by which the inflammatory response has waned) it failed to exhibit any effect on the development of fibrosis. The majority of compounds that are currently in the early phases of development rely on a wider ranging package of experimental data. There has been an important move to only use the bleomycin model therapeutically with the administration of putative antifibrotic drugs during the fibrotic phase of the model [40].

Difficulties inherent in IPF clinical trials

Expert opinion regarding best trial design and choice of end-points in IPF studies has undergone considerable change over the last decade and, as will be discussed, is currently divided [1, 2, 41–43]. Some of this division arises because of the peculiar challenges posed by IPF as a disease. Patients with IPF invariably have progressive disease that ultimately results in respiratory failure and death. The development of fibrosis in the lung causes gross architectural destruction with the loss of normal lace like structure of alveoli and thus loss of surface area for gas exchange. Even if a therapy were capable of reversing developed fibrosis in IPF it would not restore normal structure and function to the fibrotic lung. For this reason the treatment goal in IPF is to slow or, ideally, halt progression of fibrosis. As understanding of the natural history of IPF has improved it has become clear that whilst in many patients the disease progresses inexorably with a slow incremental decline in lung function and corresponding deterioration in symptoms, in a substantial minority of patients the condition remains stable for considerable periods (fig. 3). These periods of disease stability are then punctuated by rapid onset episodes of catastrophic disease progression [44]. These episodes have

been termed acute exacerbations (AEs) of IPF and are believed to represent cryptogenic flare-ups of the underlying disease process (fig. 4) [45].

The variably progressive nature of IPF creates a number of problems. First, the perfect therapy for IPF can only be anticipated to halt disease progression. Secondly, a significant proportion of patients have stable disease for prolonged periods of time, thus diluting any observed treatment effect; for this reason clinical trials in IPF require prolonged follow-up of subjects. Thirdly, the incidence of IPF requires that any adequately powered study needs a large number of recruiting centres. Fourthly, in late stage IPF complications of progressive disease, such as hypoxic respiratory failure, secondary pulmonary hypertension and pulmonary infection, frequently result in physiological deterioration or death independent of any progression in fibrosis. Fifthly, as IPF progresses, patients have difficulty performing lung function and other physiological tests, thereby, leaving trial investigators to deal with the statistical problem of missing data points. The challenge of addressing these issues has led to a number of different outcome measures being employed in late-phase IPF clinical trials (these are summarised in table 1). Observations derived from these studies have helped inform understanding of the relative merits of each of these potential primary outcomes.

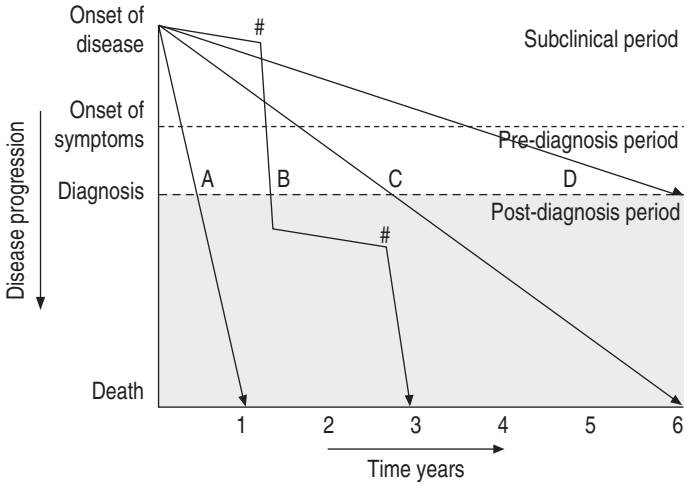


Figure 3. Schematic representation of potential clinical courses of idiopathic pulmonary fibrosis. As disease progresses, there is a subclinical period in which only radiographic findings of disease may be present, followed by a symptomatic period consisting of both pre-diagnosis and post-diagnosis clinical phases. The rate of decline and progression to death may be rapid (line A), slow (lines C and D), or mixed (line B), with periods of relative stability interposed with periods of acute decline (#). Reproduced from [44], with permission from the publisher.

Primary outcome measures

It is generally accepted that the ideal primary outcome measure for a clinical trial should provide confidence that an intervention has engaged important disease mechanisms, thus influencing disease progression with resultant change that is of meaning to an individual with that disease [1]. IPF is characterised by progressive fibroproliferation and, therefore, an effective treatment could be anticipated to halt or slow the evolution of scar tissue, thereby increasing life expectancy and reducing the burden of symptoms that might otherwise impair quality of life (QoL). Alternatively, an IPF treatment might be designed to prevent AEs, thus reducing morbidity and mortality, whilst having no effect on underlying disease progression. Other potential therapeutic options in IPF include treatment of disease complications (*e.g.* respiratory failure or pulmonary hypertension) or simply targeting symptoms (*e.g.* cough or dyspnoea). For these reasons, no single end-point is necessarily suited to all types of therapeutic intervention in IPF. Furthermore, some outcome measures are better validated than others.

Mortality

Given that IPF has a median survival of 2.8–3.5 years and 5-year outcomes worse than adenocarcinoma of the lung, mortality (which can be reported as all-cause, respiratory-related or IPF-related mortality and analysed on a time-to-event or fixed time basis) appears, at least on the face of things, to be an ideal clinical trial outcome measure for IPF. All-cause mortality is arguably

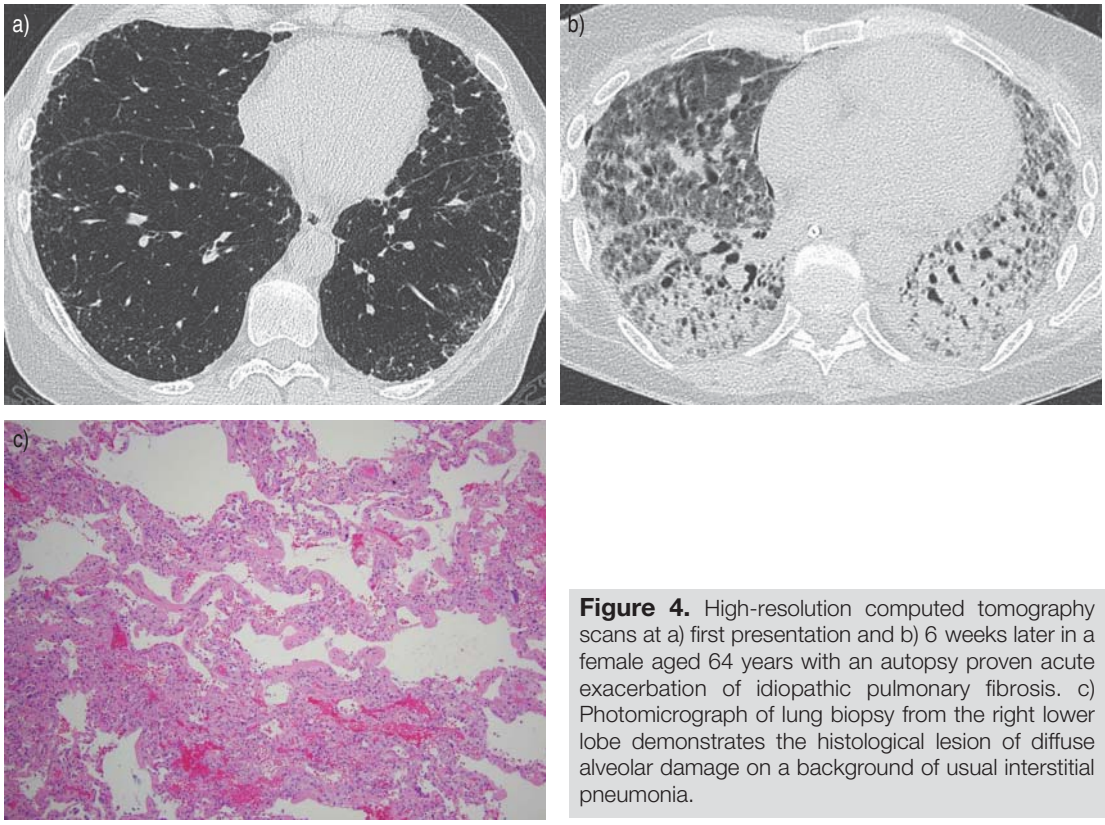


Figure 4. High-resolution computed tomography scans at a) first presentation and b) 6 weeks later in a female aged 64 years with an autopsy proven acute exacerbation of idiopathic pulmonary fibrosis. c) Photomicrograph of lung biopsy from the right lower lobe demonstrates the histological lesion of diffuse alveolar damage on a background of usual interstitial pneumonia.

the most clinically relevant of the mortality outcomes as it is well defined, reliable, easy to measure and avoids missingness or the misclassification of death by being unrelated to therapy [1]. All-cause mortality also captures off-target therapeutic effects that might increase risk of death. For these reasons, a USA-based group of IPF experts recently proposed that all-cause mortality be adopted as the primary outcome of choice for all phase III registration studies in IPF. However, there are several important reasons to challenge this viewpoint.

Feasibility

The first of these issues is feasibility. Inclusion and exclusion criteria adopted for almost all late-phase IPF trials ensure that recruited patients have mild-to-moderate disease. Individuals with advanced fibrosis (in whom it is feared that disease complications will result in mortality even if progression of fibrosis is halted), those with significant comorbidities and those with an anticipated survival of $<$12 months are usually excluded (the effects on mortality seen in the STEP-IPF study raise questions concerning the appropriateness of such exclusion criteria). Consequently, the mortality rate in IPF clinical trials tends to be considerably lower than that seen in published cohorts and epidemiology studies [32, 36, 47]. To date, only a single, phase III IPF trial, the INSPIRE study (investigating effect of interferon γ-1b on survival in patients with IPF) has adopted survival as the primary end-point [32]. In this study, in which 826 subjects were followed for a median of 64 weeks (interquartile range 41–84 weeks), the mortality rate in the placebo group was 13% and in the active treatment group was 15%. The study, which was terminated early due to lack of efficacy, was optimistically calculated to have 90% power to detect a 50% reduction in mortality with active therapy. Smaller studies, including PANTHER-IPF [49] and ACE-IPF [50] have shown a mortality signal at an earlier stage with far fewer participants, but in both cases have done so because of harm caused by the active treatment. Based on the death rates seen in the INSPIRE study and in the placebo arms of other recently published trials, it has been calculated that for an IPF clinical trial to show a statistically significant 25% reduction in mortality it would be necessary to enrol 2600 subjects for 5 years [54]. Clearly, in an era where

Table 1. Summary of late-phase idiopathic pulmonary fibrosis (IPF)

Drug	Trial acronym	Year	Study duration weeks	Subjects n	Primary end-point	Result
Interferon-γ		2004	58	330	PFS	No effect [28]
Pirfenidone		2005	36	107	Change in lowest 6MWD S_{pO_2}	Reduced acute exacerbations [27]
Warfarin		2005	57#	56	Survival time	Improved survival [30]
***N*-acetylcysteine**	IFGINIA	2005	52	182	Change in VC	Reduced progression [29]
Bosentan	BUILD-1	2008	52	58	Change in 6MWD	No effect [35]
Etanercept		2008	48	88	Change in FVC and D_{LCO}	No effect [31]
Interferon-γ	INSPIRE	2009	64#	826	Survival time	No effect [32]
Pirfenidone		2010	52	275	Change in VC	Reduced progression [34]
Imatinib		2010	96	119	Time to disease progression	No effect [33]
Sildenafil	STEP-IPF	2010	12	180	>20% increase in 6MWD	No effect [46]
Bosentan	BUILD-3	2010	52	616	Time to IPF worsening	No effect [47]
Pirfenidone	CAPACITY 1+2	2011	72	779	Change in % pred FVC	Reduced progression [36]
Nintedanib (BIBF1120)	TOMORROW	2011	52	432	Rate of FVC decline	Trend to reduced progression [48]
Prednisolone+ azathioprine	PANTHER	2012	32#	155	Change in FVC	Increased mortality [49]
Warfarin	ACE-IPF	2012	28#	145	PFS	Increased adverse events [50]
Thalidomide		2012	24	24	Cough questionnaire	Reduced cough [51]
Ambrisentan	ARTEMIS	2013	35#	492	Time to disease progression	No effect [52]
Septrin	TUPAC	2013	52	118	Change in FVC	No effect [53]

PFS: progression-free survival; 6MWD: 6-min walking distance; S_{pO_2}: arterial oxygen saturation measured by pulse oximetry; VC: vital capacity; FVC: forced vital capacity; D_{LCO}: diffusing capacity of the lung for carbon monoxide. #: median follow-up.

there are multiple ongoing global late-phase trials, this number is both unfeasible and financially ruinous even for the largest of pharmaceutical companies.

Transplantation

In many disease areas, including cancer and cardiovascular disease, it can be reasonably assumed that death is an inevitable outcome of disease progression. In IPF it remains the case that lung transplant is the one therapeutic intervention that has been shown to improve life expectancy [55]. Consequently transplantation is an option for patients with progressive disease and, although only available to a minority, it is more likely to be offered to the type of patient who is currently enrolled in a clinical trial, *i.e.* those of a younger age and with limited comorbidities. Therefore, the use of transplantation breaks the inevitable link between disease progression and death, thereby creating a statistical issue in how to handle the loss of patients to transplant in a mortality trial. Two main approaches have been taken in published studies. One is to censor individuals at the time of transplant; however, this introduces informative missingness. The second approach has been to use the date of transplant as the date of death. However, transplants are influenced by factors other than disease severity, including: donor availability and differing allocation protocols between countries and even individual hospital units, *i.e.* transplant centres, within countries. As an example of the size of the problem, in the INSPIRE study 16 subjects underwent a transplant [32].

Since then the availability of a transplant for IPF has increased in the USA due to changes in the lung allocation scoring system.

Additional therapy

Presently the majority of clinical trials on IPF are placebo controlled and exclude the use of "antifibrotic" therapy. Yet, as has been noted in Europe, Japan, Asia and Canada, pirfenidone has been licenced specifically for the treatment of IPF. Furthermore, *N*-acetylcysteine is recommended in a number of national and international guidelines as a potential therapy for IPF [56]. Participation in a placebo-controlled mortality study requires that a participant should continue in the study even as their disease progresses towards death. In doing so they would need to forego access to therapy with proven efficacy (or alternatively, the option of open-label therapy or transfer to a different clinical trial). It is clearly unethical for investigators to require this of participants who reside in countries where alternatives to no effective therapy do exist. At the time when the INSPIRE study was designed there were no realistic alternative treatments for subjects and hence the study design was acceptable at that time. Whilst the need to provide patients access to effective care does not preclude placebo controlled studies, it does require that trials are now designed to provide a contingency for individuals whose disease continues to progress. This effectively precludes mortality studies on ethical grounds, except in countries where no alternative therapies exist (or are anticipated to become available for the duration of the study).

Trial requirements

Looking to the future, it is to be hoped that more compounds will be licenced for the treatment of IPF. As this happens it will be necessary to study new therapies either in head-to-head trials or, more likely, as an add-on therapy. This will generate a series of challenges, the most important being the reduced therapeutic window in which to measure the additional benefit of new therapies under investigation. Any reduction in the anticipated rate of disease progression, and thus potentially the number of deaths, will further reduce the feasibility of using mortality as an efficacy end-point. Running trials with add-on therapies, or the comparison of multiple-drug regimens in IPF, will require careful selection of the end-points used and is likely to mandate the adaption of the currently accepted model of IPF phase III studies. Without change, trials will otherwise require ever larger numbers of participants to assess the efficacy of novel compounds.

Acute exacerbations

An AE of IPF has been defined as an unexpected, accelerated phase of symptomatic decline occurring in the absence of infection, pulmonary embolism, pneumothorax, heart failure or any other identifiable cause of acute lung injury [45]. In clinical trials, AEs have been observed in 5–8% of individuals with IPF per year [48, 57]. The in-hospital mortality associated with AEs is as high as 50% [54]. An ATS/ERS consensus group proposed the following diagnostic criteria for an AE-IPF: 1) worsening of dyspnoea within 30 days; 2) severe decrease in arterial oxygen tension; 3) the absence of pulmonary infection by endotracheal aspirate or bronchoalveolar lavage; and 4) radiographic findings of bilateral ground-glass opacities and consolidation superimposed on a pattern, typical of the UIP observed on high-frequency CT. A biopsy of AE-IPF shows the lesion of diffuse alveolar damage or organising pneumonia on a background of UIP.

The pathobiology underlying AEs is poorly understood and it is not yet known if these episodes represent an acceleration of the fibroproliferative processes characteristic of IPF or, more importantly, whether AEs are driven by a different set of mechanisms. If the latter, it is unlikely that therapies targeting disease pathways driving fibroproliferation will effectively prevent AEs. This needs to be kept in mind when considering AE frequency as an end-point for IPF trials. However, that being said, the incidence of AEs, their high mortality and the residual functional impairment, which inevitably occurs in survivors, make them an attractive treatment target. As such AE frequency has become an important potential outcome measure in IPF trials. The major key challenge around the use of AEs as an end-point surrounds the reliable identification of

AEs from other potential causes of acute decline in IPF (particularly infection). To this end, COLLARD *et al.* [58] have recently reported on the reliability of retrospective central adjudication of AEs in the STEP-IPF study (sildenafil trial of exercise performance in IPF) [58]. From the 35 serious respiratory adverse events, it was observed that four were defined as definite AEs and 14 as suspected AEs. The main difference between the two groups was the lack of bronchoscopy to exclude infection in the suspected AE group. However, overall, the outcomes were identical with a 47% mortality rate at 24 weeks. The authors concluded that it is feasible to record and adjudicate AEs in IPF studies and emphasise their importance and frequency in a clinical trial population. Whilst the challenges inherent in identifying AEs makes it a difficult measure to use as a primary end-point, AEs are of sufficient importance to be a key secondary outcome measure in late-phase IPF studies.

Disease progression

As has been noted, IPF is characterised by the progressive laying down of fibrosis. Histologically and in imaging terms this manifests as increased architectural distortion with the development of honeycomb cysts, expansion of the extracellular matrix, and the loss of alveoli. It is generally agreed that the physiological measure that best correlates with worsening fibrosis is the forced vital capacity (FVC). As has been argued cogently by DU BOIS *et al.* [42], FVC is a reliable, valid and responsive measure of disease change in IPF. The observation of reliability is based on the observed stability between screening and baseline study visits in FVC readings in several large randomised control trials. This reliability has been further enhanced in current studies by ensuring consistency of equipment, training, calibration and quality control across sites and, in some studies, ensuring central reading and verification of all spirometric flow volume loops. The importance of FVC is validated by its relationship to other key measures of disease severity including: disease extent on CTs, diffusing capacity of the lung for carbon monoxide (*D*LCO) 6-min walking distance (6MWD), oxygen saturations, symptoms and measures of QoL. FVC has been shown to be responsive to changes in fibrosis in several longitudinal studies [59–63]. In each case the change in FVC has been associated with a worse outcome. Several studies have now confirmed the importance of both 6- and 12-month categorical change in FVC. Both a 10% absolute (*e.g.* from 70% predicted down to 60% pred) and a 10% relative change (*e.g.* from 70% pred down to 63% pred) have been shown to predict a worse prognosis [63]. More recently ZAPPALLA *et al.* [59] were able to show that a categorical change, as small as a 5% relative reduction in FVC, portends a significantly worse prognosis than disease stability. However, it remains to be shown whether population change in FVC measured as a continuous variable predicts differences in survival between groups.

The simplicity of measuring FVC together with its reliability, validity and responsiveness make it a very attractive primary end-point for clinical trials. This is reflected in the number of studies that have relied on a change in FVC as their primary outcome variable (table 1). However, the use of FVC does pose a number of potential problems [64]. First, there is the issue of handling missingness. A proportion of patients with IPF will be unable to perform serial FVC measurement, either because of the development of cough, worsening respiratory failure or inter-current infection. Secondly, a proportion of patients dying in trials as a consequence of their IPF appear to have stable FVC prior to death. This is possibly because of earlier missing values or because AE is the ultimate cause of death. Imputing values for the missing data points in such individuals is frequently challenging. Thirdly, measurement variation means that even in healthy individuals with stable lung function, FVC can vary by as much as 5–9% between readings [65, 66]. This has led to a degree of confusion amongst pulmonologists in how to interpret the population changes in FVC as seen in clinical trials; in the context of an individual patient seen in the clinic. In an attempt to address this issue, DU BOIS *et al.* [62] used data derived from 1156 clinical trial participants and applied distribution based and anchor-based methods to determine the minimum clinically important difference (MCID). Using this approach the authors estimated that the MCID was 2–6% [62]. This observation requires further validation and refinement but

suggests that the aim of future IPF studies should not just be to achieve a statistically significant difference in FVC between groups but to find both a clinically important and a statistically significant difference.

Ultimately, the majority of the weaknesses for using FVC as a primary outcome measure can be overcome by a combination of careful quality control and frequent calibration of standardised equipment. The issue of missing values and a failure to detect change in individuals with terminally worsening disease can be obviated by increasing the frequency of measurement of FVC. In the early IPF clinical trials, FVC tended to be measured at baseline, 6 and 12 months [28]. This provided very limited information for the imputation of values in individuals missing the key 12-month FVC reading. By contrast, recent studies have moved to measuring FVC at least every 8 weeks, if not every 4 weeks [67]. In a 12-month study this increases the number of readings from three to 14. As such, it is possible to use the slope of change in FVC to enable a more reliable estimate for the missing values; therefore, this approach permits the use of modern statistical imputations such as bootstrapping. FVC remains, for many experts in the field, the optimal primary outcome measure for registration clinical trials [2].

Composite end-points

Composite end-points are those that combine two or more separate outcome measures. The ones that have been used in IPF trials are those of progression-free survival (PFS) and time to disease progression. PFS in the Imatinib-IPF study was defined as time to either death or 10% decline in FVC [33]. In the BUILD-3 study, IPF-worsening was defined as a decrease of FVC by 10% and a decrease in D_{LCO} of 15% or an AE or death [47]. In the ARTEMIS study, time to disease progression was defined as death, respiratory hospitalisation or categorical decline in lung function [52]. In the ACE-IPF study the time to disease progression included death, hospitalisation (non-bleeding and non-elective) or a >10% absolute decline in FVC [50].

PFS is a concept that first emerged in oncology trials [68]. PFS appears, at least superficially, to both overcome the problem of incorporating the different aspects of disease that are impacted by disease progression and minimises the effect of missing data points caused by death. However, as previously illustrated earlier on, there is a lack of consensus surrounding how best to define PFS in IPF, resulting in definitions that vary from study to study. In basic terms the aim of any new therapy in IPF is either to improve the QoL or quantity of life (or ideally both) for an individual with the disease. The same is true of studies undertaken in advanced cancer. However, in the field of oncology, PFS has been criticised for a number of reasons, most importantly of which is that it does not mirror either improvements in overall survival or QoL [68]. The lack of correlation with changes in QoL occurs for the simple reason that there are studies lacking that can map the effect of PFS (or the individual components of PFS) to changes in QoL; this same criticism applies to the use of PFS in IPF. The failure of PFS to correlate with overall survival in oncology is more challenging to explain. First, there is a requirement to undertake the necessary validation studies to correlate PFS with overall survivals. The second issue is that unlike death, which occurs at a given and clearly identifiable time, many of the components of PFS occur at an unknown time and are only detected at the time of predetermined study interventions (*e.g.* in the context of IPF a 10% decline in FVC can only be measured using spirometry, so to ascertain the decline and the relationship between the date on which it is measured compared with the date on which it actually occurred, the frequency with which the spirometry is undertaken will influence the outcome). Thirdly, is the possibility that changes in one element of PFS might result in biological changes that alter disease behaviour and, thus, influence longer term survival, *e.g.* put in an IPF context it could be possible to slow disease progression but increase frequency of AEs and thus have a neutral effect on overall survival.

None of these issues precludes the use of composite measurements as clinical trial end-points in IPF. However, each issue speaks to a need to ensure that chosen end-points are appropriately longitudinally validated. At present the choice of components for composite end-points are driven

by observations derived from small, often single centre, retrospective cohort studies. The analysis of combined study cohorts from completed IPF clinical trials has provided some insights in to the relative merits of individual aspects of disease progression [61, 62]. The cooperative use of these valuable datasets will be important if a robust composite end-point is to be identified. Importantly any composite outcome measure must be shown to provide information that is additional to that provided by key individual components, *e.g.* FVC.

Changes in functional status

In the first Japanese pirfenidone study [27] and in BUILD-1 [35], components of the 6-min walk test (6MWT) were used as the primary end-point. The 6MWT seeks to provide a simple and understandable measurement of an individual's functional status, *i.e.* how much their disease impacts on their ability to undertake activities of daily living such as walking to the shops. Both baseline 6MWD and maximal desaturation occurring during a 6MWT have been correlated with subsequent survival [69]. A change in the 6MWD over time has also been correlated with prognosis and a MCID, at a population level, in 6MWD defined as 28 m [70, 71].

Whilst the cardinal feature of IPF is exercise limiting dyspnoea, submaximal exercise testing is influenced by a number of factors other than just the diffusion defect caused by the development and evolution of fibrosis. These limiting factors include pulmonary hypertension, cardiovascular disease, joint or locomotor disease, general fitness levels and, importantly, motivation. Added to this there is the practical challenge of standardising 6MWTs between units, not all of which will have sufficiently long, straight, pedestrian free corridors. Also, in a disease characterised by a diffusion defect, the difference in atmospheric oxygen concentration between sea level and altitude can introduce an important confounder in multicentre studies. For these reasons there is a high degree of variability in 6MWD even when repeated at short time intervals in motivated patients. This variability combined with the issue of missing data (usually from patients with progressive disease who are no longer able to undertake 6MWTs) significantly impacts the utility of 6MWD or other 6MWT variables as primary outcome measures in IPF studies.

Secondary outcome measures

To date a wide range of physiological, imaging, safety and patient reported outcome measures have been used as secondary end-points in clinical trials. The evidence to support the use of these outcomes is, in many cases, limited and too often has added little information over and above that provided by the primary end-point. Nonetheless, these other potential IPF trial outcomes will be considered in more detail.

Physiology

As well as 6MWD and FVC other measures of pulmonary physiology have been linked to outcomes in IPF. A change in *D*LCO has, like FVC, been shown to predict a worse prognosis when analysed categorically (with a 15% change being the important threshold) [60]. Unlike FVC, *D*LCO also reflects change in the pulmonary vasculature and so drops with the development of pulmonary hypertension. The use of *D*LCO as a study outcome is hampered by inherent variability of the test. *D*LCO is harder than FVC to standardise and provide a quality control across centres. For this reason *D*LCO has not been used as a standalone primary outcome measure in IPF studies but has been used as a component of PFS. Other physiological measures that have been utilised include a change in the alveolar–arterial oxygen tension difference and change in resting oxygen saturation [28]. Neither the reliability nor the reproducibility of either measure has been prospectively validated. Furthermore, arterial oxygen levels can be influenced by activity; especially in individuals on the steep part of the oxygen dissociation curve. For this reason standardisation of measurements is difficult, limiting the use of oxygen as a major study outcome measure.

Biomarkers

A biomarker is generally accepted to be a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacological responses to therapeutic interventions [72]. A biomarker of treatment response would afford the potential to simplify and shorten IPF clinical trials and would enable the development of robust proof-of-concept trials. Although a number of studies have been conducted to identify biomarkers in IPF, almost all of these have been opportunistic and are hampered by small sample size, retrospective design, a lack of validation, and an absence of longitudinal sampling. Nonetheless, a number of potential markers have been identified. Markers associated with worse patient survival include: elevated levels of mucin-1 (MUC1/KL6) [73], CC chemokine ligand (CCL-18) [74], surfactant protein A [75], circulating fibrocytes [76], periostin [77] and semaphoring-7a positive regulatory T cells [78]. In the biggest published biomarker study to date, RICHARDS *et al.* [79] analysed 92 proteins in the blood of 241 patients with IPF. Dividing there subjects into derivation and validation cohorts, the authors were able to show that high concentrations of matrix metalloproteinase-7, intercellular adhesion molecule-1 and interleukin-8 predicted poor survival. Whilst in their totality these results strongly suggest that measurements of serum markers may provide information about disease activity in IPF, there is considerable further research needed before any of these putative biomarkers can be used, either to stratify or measure outcome, in IPF clinical trials.

Patient-reported outcome measures

Patient-reported outcome (PRO) measures are those which seek to ''directly report the status of a patient's health condition without the interpretation of the patient's response by a clinician or anybody else'' [1]. PROs can seek to report the impact of specific symptoms (*e.g.* cough or dyspnoea), or can be disease or organ specific or else provide information about broader constructs *e.g.* health status. While a range of different PROs exist for respiratory disease, very few have been validated for IPF and even fewer have been followed longitudinally to test responsiveness to disease progression. One of the most commonly used respiratory disease PROs, the St George's Respiratory Questionnaire (SGRQ) was originally developed for use in chronic obstructive pulmonary disease but has recently been validated in IPF [80]. Similarly the 36-item short-form health survey (SF-36) questionnaire has also been validated in IPF [81]. SWIGRIS *et al.* [81] have estimated that the MCID in SF-36 in IPF is 2–4 points whilst for the SGRQ it is 5–8 points. A number of disease specific IPF QoL instruments have been developed but require further research before they can be implemented as research tools. The University California San Diego shortness of breath questionnaire has, in the recent STEP-IPF study, been shown to have validity as a measure of change in dyspnoea over time inIPF patients [82]. Similarly, in a recent study of thalidomide for cough in IPF, the Cough Quality of Life Questionnaire has been shown to be valid and reliable as a symptom measure in IPF [83].

The recent introduction of pirfenidone across Europe has highlighted that healthcare agencies (responsible for the licensing and regulation of drugs) and healthcare funders (responsible for reimbursement) have different expectations from clinical trial data. Whilst regulators expect to see clear efficacy and safety signals, funders place a high premium on meaningful benefit for patients. As such PROs are likely to become established as important secondary end-point measure in future IPF trials.

Imaging

HRCT imaging has revolutionised the diagnosis of IPF and the CT appearance is an important determinant of an individual patient's suitability for clinical trial involvement. At baseline CT

provides an assessment of disease extent that correlates with physiological measures of severity [84]. However, attempts to use a change in CT for the extent of the disease as a trial end-point have been hampered by inter-individual variability in reporting and a lack of sensitivity for detecting subtle changes (in part this is magnified by the challenge of ensuring that serial CT scans are technically identical with similar inspiratory effort). Computerised reporting of co-registered scans may help overcome some of the limitations of detecting change in disease extent on CT. RAGHU *et al.* [85], in a recent proof-of-concept study with the anti-connective tissue growth-factor antibody FG-3019, reported the use of quantified HRCT scores, which the authors used as a primary study end-point. Whilst of interest, this approach has yet to be longitudinally validated.

Another imaging modality that has potential for use as a trial end-point is positron emission tomography (PET). WIN *et al.* [86] recently reported on the reproducibility of 2-fluoro-2-deoxy-D-glucose (FDG)-PET signalling in the lungs of individuals with IPF. The group postulated on its use as a trial end-point. In animal models of fibrosis bespoke PET ligands have been used to demonstrate engagement of target mechanisms by novel therapeutic compounds. The use of PET imaging may, therefore, have a role to play in early proof-of-concept or proof-of-mechanism in IPF studies. However, the burden of radiation that accompanies PET scans is too high for their routine use in later phase clinical trials.

Early-phase clinical trials

The majority of studies undertaken in IPF to date have been opportunistic; that is they have been carried out with compounds generally developed for other indications rather than designed specifically as antifibrotic agents (pirfenidone can be considered the one exception to this phenomenon). In most cases these drugs have been tested directly in phase IIB or even phase III trials. Whilst there has been a clear evolution in late-phase clinical trial design in IPF, with improved understanding of the relative value of the different end-points measured, the same cannot be said for proof-of-concept or proof-of-mechanism studies. Appropriate, validated, short-term readouts that provide confidence that a novel therapy has engaged both its target and is blocking key fibrotic pathways are lacking. Until such markers are available it will be difficult to design short, early-phase and dose-validation studies for IPF. Given the large number of potential targets in IPF there is a major need for proof-of-concept studies as a means of narrowing down candidate target and molecule selection and to expedite the passage of the most promising drugs from early clinical development through to registration studies and then the clinic. Exploratory omics-based research in populations of carefully phenotyped patients with IPF will hopefully lead to the identification of candidate outcome measures. Without such measures the drug development pipeline in IPF remains incomplete [87].

Conclusion

The clinical development and successful licensing in Europe, Asia and Japan of pirfenidone has raised expectations for individuals suffering with IPF and their clinicians. Once considered a disease too rare and difficult to treat to merit the development of clinical trials it has now become clear that IPF represents a major healthcare burden in the developed world. With the arrival of clinical trials has come improved understanding of disease, including the recognition of the frequency and seriousness of AEs. Clinical trial end-points that were chosen in early studies have subsequently been abandoned as their limitations have become clear. Regulators, patient advocacy groups and trial clinicians have worked hard to identify meaningful outcome measures for use in late-phase clinical trials. Notwithstanding recent debates, the design of registration studies for IPF is reaching maturity. Furthermore, regulatory agencies have clearly articulated the need for late-phase trials to achieve their primary end-point whilst also

demonstrating changes in secondary outcomes, which also support efficacy. However, considerable further enabling research is required to identify optimal outcome measures for use in small early-phase studies so as to permit the development of a more efficient clinical trial pipeline. Given advances in understanding clinical trial design and choice of end-point, which have arisen in the last decade, it is to be hoped that this major remaining hurdle can be overcome soon. At the same time further work is required to relate changes in specific IPF outcomes to change in parameters that are considered clinically meaningful by sufferers of this devastating disease.

Statement of Interest

T.M. Maher is in receipt of an unrestricted academic industry grant from GSK. In the last 3 years he has received advisory board or consultancy fees from Actelion, Boehringer Ingelheim, GSK, Respironics, InterMune and Sanofi-Aventis. T.M. Maher has received speaker's fees from UCB, Boehringer Ingelheim, InterMune and AstraZeneca and his institution has received an unrestricted educational grant from InterMune and consultancy fees on his behalf from Novartis and Takeda.

References

1. Raghu G, Collard HR, Anstrom KJ, *et al.* Idiopathic pulmonary fibrosis: clinically meaningful primary endpoints in phase 3 clinical trials. *Am J Respir Crit Care Med* 2012; 185: 1044–1048.
2. Wells AU, Behr J, Costabel U, *et al.* Hot of the breath: mortality as a primary end-point in IPF treatment trials: the best is the enemy of the good. *Thorax* 2012; 67: 938–940.
3. Richeldi L. Idiopathic pulmonary fibrosis: current challenges and future perspectives. *Eur Respir Rev* 2013; 22: 103–105.
4. Maher TM. Pirfenidone in idiopathic pulmonary fibrosis. *Drugs Today (Barc.)* 2010; 46: 473–482.
5. Maher TM, Wells AU, Laurent GJ. Idiopathic pulmonary fibrosis: multiple causes and multiple mechanisms? *Eur Respir J* 2007; 30: 835–839.
6. Gribbin J, Hubbard RB, Le JI, *et al.* Incidence and mortality of idiopathic pulmonary fibrosis and sarcoidosis in the UK. *Thorax* 2006; 61: 980–985.
7. Monaghan H, Wells AU, Colby TV, *et al.* Prognostic implications of histologic patterns in multiple surgical lung biopsies from patients with idiopathic interstitial pneumonias. *Chest* 2004; 125: 522–526.
8. Bjoraker JA, Ryu JH, Edwin MK, *et al.* Prognostic significance of histopathologic subsets in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 1998; 157: 199–203.
9. Navaratnam V, Fleming KM, West J, *et al.* The rising incidence of idiopathic pulmonary fibrosis in the U.K. *Thorax* 2011; 66: 462–467.
10. Raghu G, Weycker D, Edelsberg J, *et al.* Incidence and prevalence of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2006; 174: 810–816.
11. Lee JS, Collard HR, Raghu G, *et al.* Does chronic microaspiration cause idiopathic pulmonary fibrosis? *Am J Med* 2010; 123: 304–311.
12. Stewart JP, Egan JJ, Ross AJ, *et al.* The detection of Epstein-Barr virus DNA in lung tissue from patients with idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 1999; 159: 1336–1341.
13. Marshall RP, Puddicombe A, Cookson WO, *et al.* Adult familial cryptogenic fibrosing alveolitis in the United Kingdom. *Thorax* 2000; 55: 143–146.
14. Nogee LM, Dunbar AE 3rd, Wert SE, *et al.* A mutation in the surfactant protein C gene associated with familial interstitial lung disease. *N Engl J Med* 2001; 344: 573–579.
15. Seibold MA, Wise AL, Speer MC, *et al.* A common MUC5B promoter polymorphism and pulmonary fibrosis. *N Engl J Med* 2011; 364: 1503–1512.
16. Tsakiri KD, Cronkhite JT, Kuan PJ, *et al.* Adult-onset pulmonary fibrosis caused by mutations in telomerase. *Proc Natl Acad Sci USA* 2007; 104: 7552–7557.
17. Fingerlin TE, Murphy E, Zhang W, *et al.* Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. *Nat Genet* 2013; 45: 613–620.
18. Noth I, Zhang Y, Ma SF, *et al.* Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: a genome-wide association study. *Lancet Respir Med* 2013; 1: 309–317.
19. Maher TM. Idiopathic pulmonary fibrosis: pathobiology of novel approaches to treatment. *Clin Chest Med* 2012; 33: 69–83.
20. American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the idiopathic interstitial pneumonias. *Am J Respir Crit Care Med* 2002; 165: 277–304.
21. Raghu G, Collard HR, Egan JJ, *et al.* An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med* 2011; 183: 788–824.
22. Maher TM. A clinical approach to diffuse parenchymal lung disease. *Immunol Allergy Clin North Am* 2012; 32: 453–472.

23. Johnson MA, Kwan S, Snell NJ, *et al.* Randomised controlled trial comparing prednisolone alone with cyclophosphamide and low dose prednisolone in combination in cryptogenic fibrosing alveolitis. *Thorax* 1989; 44: 280–288.
24. Raghu G, Depaso WJ, Cain K, *et al.* Azathioprine combined with prednisone in the treatment of idiopathic pulmonary fibrosis: a prospective double-blind, randomized, placebo-controlled clinical trial. *Am Rev Respir Dis* 1991; 144: 291–296.
25. Douglas WW, Ryu JH, Swensen SJ, *et al.* Colchicine versus prednisone in the treatment of idiopathic pulmonary fibrosis. A randomized prospective study. *Am J Respir Crit Care Med* 1998; 158: 220–225.
26. Ziesche R, Hofbauer E, Wittmann K, *et al.* A preliminary study of long-term treatment with interferon γ-1b and low-dose prednisolone in patients with idiopathic pulmonary fibrosis. *N Engl J Med* 1999; 341: 1264–1269.
27. Azuma A, Nukiwa T, Tsuboi E, *et al.* Double-blind, placebo-controlled trial of pirfenidone in patients with idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2005; 171: 1040–1047.
28. Raghu G, Brown KK, Bradford WZ, *et al.* A placebo-controlled trial of interferon γ-1b in patients with idiopathic pulmonary fibrosis. *N Engl J Med* 2004; 350: 125–133.
29. Demedts M, Behr J, Buhl R, *et al.* High-dose acetylcysteine in idiopathic pulmonary fibrosis. *N Engl J Med* 2005; 353: 2229–2242.
30. Kubo H, Nakayama K, Yanai M, *et al.* Anticoagulant therapy for idiopathic pulmonary fibrosis. *Chest* 2005; 128: 1475–1482.
31. Raghu G, Brown KK, Costabel U, *et al.* Treatment of idiopathic pulmonary fibrosis with etanercept: an exploratory, placebo-controlled trial. *Am J Respir Crit Care Med* 2008; 178: 948–955.
32. King TE Jr, Albera C, Bradford WZ, *et al.* Effect of interferon γ-1b on survival in patients with idiopathic pulmonary fibrosis (INSPIRE): a multicentre, randomised, placebo-controlled trial. *Lancet* 2009; 374: 222–228.
33. Daniels CE, Lasky JA, Limper AH, *et al.* Imatinib treatment for idiopathic pulmonary fibrosis: randomized placebo-controlled trial results. *Am J Respir Crit Care Med* 2010; 181: 604–610.
34. Taniguchi H, Ebina M, Kondoh Y, *et al.* Pirfenidone in idiopathic pulmonary fibrosis. *Eur Respir J* 2010; 35: 821–829.
35. King TE Jr, Behr J, Brown KK, *et al.* BUILD-1: a randomized placebo-controlled trial of bosentan in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2008; 177: 75–81.
36. Noble PW, Albera C, Bradford WZ, *et al.* Pirfenidone in patients with idiopathic pulmonary fibrosis (CAPACITY): two randomised trials. *Lancet* 2011; 377: 1760–1769.
37. du Bois RM. Strategies for treating idiopathic pulmonary fibrosis. *Nat Rev Drug Discov* 2010; 9: 129–140.
38. Moeller A, Ask K, Warburton D, *et al.* The bleomycin animal model: a useful tool to investigate treatment options for idiopathic pulmonary fibrosis? *Int J Biochem Cell Biol* 2008; 40: 362–382.
39. Daniels CE, Wilkes MC, Edens M, *et al.* Imatinib mesylate inhibits the profibrogenic activity of TGF-β and prevents bleomycin-mediated lung fibrosis. *J Clin Invest* 2004; 114: 1308–1316.
40. Maher TM, Wells AU. Lost in translation; from animal models of pulmonary fibrosis to human disease. *Respirology* 2009; 14: 915–916.
41. Vancheri C, du Bois RM. A progression-free end-point for idiopathic pulmonary fibrosis trials: lessons from cancer. *Eur Respir J* 2013; 41: 262–269.
42. du Bois RM, Nathan SD, Richeldi L, *et al.* Idiopathic pulmonary fibrosis: lung function is a clinically meaningful endpoint for phase III trials. *Am J Respir Crit Care Med* 2012; 186: 712–715.
43. Spagnolo P, Luppi F, Maher TM, *et al.* Primary endpoints in phase 3 clinical trials in idiopathic pulmonary fibrosis: one step at a time. *Am J Respir Crit Care Med* 2013; 187: 1271–1272.
44. Ley B, Collard HR, King TE Jr. Clinical course and prediction of survival in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2011; 183: 431–440.
45. Collard HR, Moore BB, Flaherty KR, *et al.* Acute exacerbations of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2007; 176: 636–643.
46. Idiopathic Pulmonary Fibrosis Clinical Research Network, Zisman DA, Schwarz M, *et al.* A controlled trial of sildenafil in advanced idiopathic pulmonary fibrosis. *N Engl J Med* 2010; 363: 620–628.
47. King TE Jr, Brown KK, Raghu G, *et al.* BUILD-3: a randomized, controlled trial of bosentan in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2011; 184: 92–99.
48. Richeldi L, Costabel U, Selman M, *et al.* Efficacy of a tyrosine kinase inhibitor in idiopathic pulmonary fibrosis. *N Engl J Med* 2011; 365: 1079–1087.
49. Idiopathic Pulmonary Fibrosis Clinical Research Network. Prednisone, azathioprine, and *N*-acetylcysteine for pulmonary fibrosis. *N Engl J Med* 2012; 366: 1968–1977.
50. Noth I, Anstrom KJ, Calvert SB, *et al.* A placebo-controlled randomized trial of warfarin in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2012; 186: 88–95.
51. Horton MR, Santopietro V, Mathew L, *et al.* Thalidomide for the treatment of cough in idiopathic pulmonary fibrosis: a randomized trial. *Ann Intern Med* 2012; 157: 398–406.
52. Raghu G, Behr J, Brown KK, *et al.* Treatment of idiopathic pulmonary fibrosis with ambrisentan: a parallel, randomized trial. *Ann Intern Med* 2013; 158: 641–649.
53. Shulgina L, Cahn AP, Chilvers ER, *et al.* Treating idiopathic pulmonary fibrosis with the addition of co-trimoxazole: a randomised controlled trial. *Thorax* 2013; 68: 155–162.

54. Bradford WZ, Cohen AH, Leff JA. Selection of clinically meaningful primary endpoints in phase 3 clinical trials in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2013; 187: 1269–1270.
55. Keating D, Levvey B, Kotsimbos T, *et al.* Lung transplantation in pulmonary fibrosis: challenging early outcomes counterbalanced by surprisingly good outcomes beyond 15 years. *Transplant Proc* 2009; 41: 289–291.
56. Bradley B, Branley HM, Egan JJ, *et al.* Interstitial lung disease guideline: the British Thoracic Society in collaboration with the Thoracic Society of Australia and New Zealand and the Irish Thoracic Society. *Thorax* 2008; 63: Suppl. 5, v1–v58.
57. Martinez FJ, Safrin S, Weycker D, *et al.* The clinical course of patients with idiopathic pulmonary fibrosis. *Ann Intern Med* 2005; 142: 963–967.
58. Collard HR, Yow E, Richeldi L, *et al.* Suspected acute exacerbation of idiopathic pulmonary fibrosis as an outcome measure in clinical trials. *Respir Res* 2013; 14: 73.
59. Zappala CJ, Latsi PI, Nicholson AG, *et al.* Marginal decline in forced vital capacity is associated with a poor outcome in idiopathic pulmonary fibrosis. *Eur Respir J* 2010; 35: 830–836.
60. Latsi PI, du Bois RM, Nicholson AG, *et al.* Fibrotic idiopathic interstitial pneumonia: the prognostic value of longitudinal functional trends. *Am J Respir Crit Care Med* 2003; 168: 531–537.
61. du Bois RM, Weycker D, Albera C, *et al.* Ascertainment of individual risk of mortality for patients with idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2011; 184: 459–466.
62. du Bois RM, Weycker D, Albera C, *et al.* Forced vital capacity in patients with idiopathic pulmonary fibrosis: test properties and minimal clinically important difference. *Am J Respir Crit Care Med* 2011; 184: 1382–1389.
63. Richeldi L, Ryerson CJ, Lee JS, *et al.* Relative *versus* absolute change in forced vital capacity in idiopathic pulmonary fibrosis. *Thorax* 2012; 67: 407–411.
64. Collard HR. Idiopathic pulmonary fibrosis and pirfenidone. *Eur Respir J* 2010; 35: 728–729.
65. Dowson LJ, Mushtaq M, Watts T, *et al.* A re-audit of pulmonary function laboratories in the West Midlands. *Respir Med* 1998; 92: 1155–1162.
66. Tantucci C, Pinelli V, Cossi S, *et al.* Reference values and repeatability of inspiratory capacity for men and women aged 65-85. *Respir Med* 2006; 100: 871–877.
67. Richeldi L, Brown KK, Costabel U, *et al.* Efficacy of the tyrosine kinase inhibitor BIBF 1120 in patients with IPF: consistent pattern of primary endpoint results in sensitivity analyses of the TOMORROW Trial. *Am J Respir Crit Care Med* 2012; 185: A3633.
68. Booth CM, Eisenhauer EA. Progression-free survival: meaningful or simply measurable? *J Clin Oncol* 2012; 30: 1030–1033.
69. Flaherty KR, Andrei AC, Murray S, *et al.* Idiopathic pulmonary fibrosis: prognostic value of changes in physiology and six-minute-walk test. *Am J Respir Crit Care Med* 2006; 174: 803–809.
70. du Bois RM, Weycker D, Albera C, *et al.* Six-minute-walk test in idiopathic pulmonary fibrosis: test validation and minimal clinically important difference. *Am J Respir Crit Care Med* 2011; 183: 1231–1237.
71. Swigris JJ, Wamboldt FS, Behr J, *et al.* The 6 minute walk in idiopathic pulmonary fibrosis: longitudinal changes and minimum important difference. *Thorax* 2010; 65: 173–177.
72. Jones PW, Agusti AGN. Outcomes and markers in the assessment of chronic obstructive pulmonary disease. *Eur Respir J* 2006; 27: 822–832.
73. Yokoyama A, Kondo K, Nakajima M, *et al.* Prognostic value of circulating KL-6 in idiopathic pulmonary fibrosis. *Respirology* 2006; 11: 164–168.
74. Prasse A, Probst C, Bargagli E, *et al.* Serum CC-chemokine ligand 18 concentration predicts outcome in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2009; 179: 717–723.
75. Greene KE, King TE Jr, Kuroki Y, *et al.* Serum surfactant proteins-A and -D as biomarkers in idiopathic pulmonary fibrosis. *Eur Respir J* 2002; 19: 439–446.
76. Moeller A, Gilpin SE, Ask K, *et al.* Circulating fibrocytes are an indicator of poor prognosis in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2009; 179: 588–594.
77. Naik PK, Bozyk PD, Bentley JK, *et al.* Periostin promotes fibrosis and predicts progression in patients with idiopathic pulmonary fibrosis. *Am J Physiol Lung Cell Mol Physiol* 2012; 303: L1046–L1056.
78. Reilkoff RA, Peng H, Murray LA, *et al.* Semaphorin 7a+ regulatory T cells are associated with progressive idiopathic pulmonary fibrosis and are implicated in transforming growth factor-β1-induced pulmonary fibrosis. *Am J Respir Crit Care Med* 2013; 187: 180–188.
79. Richards TJ, Kaminski N, Baribaud F, *et al.* Peripheral blood proteins predict mortality in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2012; 185: 67–76.
80. Yorke J, Jones PW, Swigris JJ. Development and validity testing of an IPF-specific version of the St George's Respiratory Questionnaire. *Thorax* 2010; 65: 921–926.
81. Swigris JJ, Brown KK, Behr J, *et al.* The SF-36 and SGRQ: validity and first look at minimum important differences in IPF. *Respir Med* 2010; 104: 296–304.
82. Swigris JJ, Han M, Vij R, *et al.* The UCSD shortness of breath questionnaire has longitudinal construct validity in idiopathic pulmonary fibrosis. *Respir Med* 2012; 106: 1447–1455.
83. Lechtzin N, Hilliard ME, Horton MR. Validation of the Cough Quality-of-Life Questionnaire in patients with idiopathic pulmonary fibrosis. *Chest* 2013; 143: 1745–1749.

84. Wells AU, Desai SR, Rubens MB, *et al.* Idiopathic pulmonary fibrosis: a composite physiologic index derived from disease extent observed by computed tomography. *Am J Respir Crit Care Med* 2003; 167: 962–969.
85. Raghu G, Scholand MB, de Andrade J, *et al.* Phase 2 trial of FG-3019, anti-CTGF monoclonal antibody, in idiopathic pulmonary fibrosis (IPF): Preliminary safety and efficacy results. *Eur Respir J* 2012; 40: Suppl. 56, 511s.
86. Win T, Lambrou T, Hutton BF, *et al.* 18F-Fluorodeoxyglucose positron emission tomography pulmonary imaging in idiopathic pulmonary fibrosis is reproducible: implications for future clinical trials. *Eur J Nucl Med Mol Imaging* 2012; 39: 521–528.
87. Maher TM. PROFILEing idiopathic pulmonary fibrosis: rethinking biomarker discovery. *Eur Respir Rev* 2013; 22: 148–152.

Chapter 4

Cystic fibrosis

Reshma Amin[*] *and Felix Ratjen*[#]

SUMMARY: The most commonly used outcome measures in clinical trials in cystic fibrosis (CF) patients include forced expiratory volume in 1 s (FEV1), pulmonary exacerbations and, more recently, health-related quality of life. These three outcomes are considered surrogate end-points by the CF research community and regulatory agencies, as improvements in these parameters are likely to predict clinical benefit. However, with improved patient clinical status, pulmonary function tests such as FEV1 are often within the normal range and pulmonary exacerbations requiring hospitalisation have become a less frequent event. Hence, there is a need for new, more sensitive outcome measures. Novel end-points undergo a rigorous validation process before they can be used as primary outcomes in pivotal clinical trials. Here, we review the established surrogate end-points used in clinical trials in CF and provide an overview of other promising end-points, such as multiple-breath washout technology, imaging technology, measures of airway infection and inflammation focusing on bronchoalveolar lavage and sputum, as well as intestinal current measurements.

[*]Dept of Respiratory Medicine, The Hospital for Sick Children, Toronto, and
[#]Dept of Paediatric Respiratory Medicine, The Hospital for Sick Children, Toronto, ON, Canada.

Correspondence: R. Amin, Dept of Respiratory Medicine, The Hospital for Sick Children, 555 University Avenue, Toronto, ON, M5G 1X8, Canada.
Email: reshma.amin@sickkids.ca

Eur Respir Monogr 2013; 62: 54–69.
Copyright ERS 2013.
DOI: 10.1183/1025448x.10044612
Print ISBN: 978-1-84984-044-6
Online ISBN: 978-1-84984-045-3
Print ISSN: 1025-448x
Online ISSN: 2075-6674

Cystic fibrosis (CF) is one of the most prevalent fatal, autosomal recessive diseases in Caucasians, occurring in one in 2500 live births [1, 2]. CF is the result of a defect in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene, which is expressed in the epithelium of several organs, including the lungs, pancreas, gastrointestinal tract, reproductive tract, skin and nasal mucosa. The lack of functional CFTR protein results in disease manifestations in organs where CFTR has relevant physiological roles. However, the pulmonary manifestations account for over 90% of the morbidity and mortality in CF patients. Although the temporal relationship between infection and inflammation remains unclear in early disease, there is an ongoing cycle of airway obstruction, chronic bacterial infection and inflammation, leading to structural damage to the airway [3, 4]. Bronchiectasis develops, which eventually leads to respiratory failure and premature death [5].

Improvements in the life expectancy of CF patients have been driven by the development of new drugs evaluated in interventional trials and by implementation of these therapies in clinical care. The most commonly used outcome measures in clinical trials in CF patients include forced expiratory volume in 1 s (FEV1), pulmonary exacerbations and, more recently, the patient-reported outcome (PRO) of health-related quality of life (HRQoL). These three outcomes are considered established surrogate end-points by the CF research community and regulatory agencies, as improvements in these parameters are likely to predict clinical benefit [6]. However, with the introduction of newborn screening and improved patient

clinical status, pulmonary function tests such as FEV_1 are often within the normal range and pulmonary exacerbations requiring hospitalisation have become a less frequent event. Hence, there is a need for new, more sensitive outcome measures. However, these end-points must undergo a rigorous validation process before they can be used as primary outcomes in pivotal clinical trials. Here, we review the established surrogate end-points used in clinical trials in CF and provide an overview of other promising end-points, such as multiple-breath washout technology, imaging technology, measures of airway infection and inflammation focusing on bronchoalveolar lavage (BAL) and sputum, as well as intestinal current measurements (ICMs).

Established surrogate end-points

Spirometry

Spirometry provides physiological information about lung function through a series of breathing manoeuvres. In general, FEV_1, forced vital capacity (FVC) and forced expiratory flow at 25–75% of FVC ($FEF_{25–75\%}$) are captured and monitored in clinical care and interventional studies. However, FEV_1 is the most commonly used parameter in research trials as well as for the clinical management of CF patients, as it has been shown to be a strong predictor of morbidity and mortality [7, 8]. Although FEV_1 had long been assumed to be associated with mortality, COREY and FAREWELL [7] published a landmark paper that demonstrated this relationship. They reviewed the largest existing database of 3795 CF patients over a 20-year period and found that of several clinical variables, including body mass, and the presence of *Pseudomonas aeruginosa* and *Burkholderia cepacia* in the sputum, FEV_1 was the strongest predictor of mortality. The analysis demonstrated that, at any level of FEV_1 measured cross-sectionally, FEV_1 predicts subsequent decline and mortality. A drop in FEV_1 of 10% predicted corresponded to a doubling of the hazard ratio and risk of death. COREY *et al.* [8] went on to show the average rate of decline of spirometry measurements in subgroups on the basis of survival age, sex, pancreatic status and genotype in 366 CF patients (fig. 1). As such, FEV_1 has become the main outcome measure used in therapeutic intervention studies as well as in the daily clinical management of CF patients.

In the last decade, it has been established that CF lung disease begins shortly after birth, and progresses despite a lack of clinical signs and symptoms [9–12]. As a result, there is an increasing emphasis on early intervention for CF lung disease [13]. FEV_1 is of limited use in CF patients with mild lung disease because of the insensitivity to peripheral airway disease. $FEF_{25–75\%}$ is sensitive to small airway disease but its coefficient of variation is higher than that of FEV_1 [14]; therefore, the variability between measurements is a limiting factor for its use as a clinical end-point.

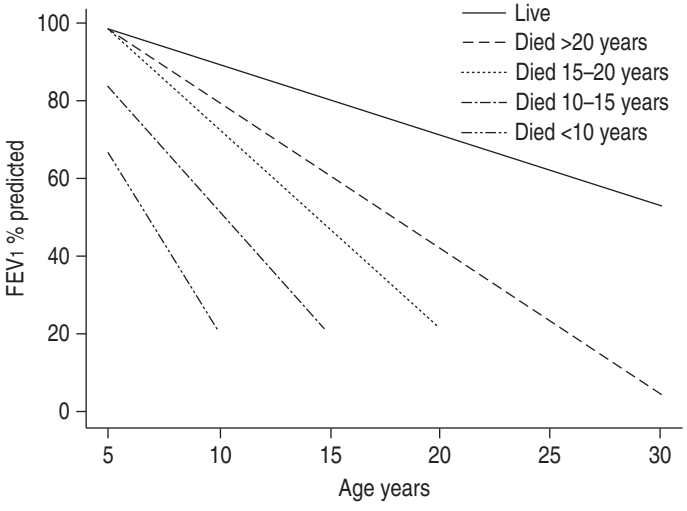


Figure 1. Mixed model regression lines of forced expiratory volume in 1 s (FEV_1) *versus* age for five retrospectively determined disease severity categories according to age at time of death. Reproduced from [8] with permission from the publisher.

Other limitations of FEV_1 % pred include its relative stability and the need for active cooperation to perform the manoeuvre. Therefore,

children with CF can have significant pulmonary disease as demonstrated by computed tomography (CT) imaging while maintaining an FEV1 greater than 80% pred [15, 16]. Furthermore, clinical trials in children powered for a change in FEV1 require large sample sizes or need to follow children over prolonged periods of time because of the relative stability of FEV1 in those with mild disease. Finally, spirometry cannot be reliably performed in children less than 6 years of age because active cooperation is required to perform the manoeuvre [17–20]. Therefore, more sensitive measures of lung function are needed for clinical studies in CF patients with early, mild disease.

Pulmonary exacerbations

A pulmonary exacerbation is defined as a clinically relevant event for patients with CF in which there is a progressive decline in FEV1 accompanied by acute worsening of clinical symptoms, such as increased cough, increased sputum production, dyspnoea, anorexia and weight loss [21, 22]. Exacerbations have been shown to increase with age as well as with more severe lung function impairment [21]. Pulmonary exacerbations are a useful established surrogate outcome measure in clinical trials in CF because of their association with survival [23, 24]. Furthermore, pulmonary exacerbations are also a risk factor for the rate of decline of FEV1 [25]. Recently, the effect of pulmonary exacerbations on the long-term evolution of lung function has also been reported [26]. In this retrospective cohort study, a pulmonary exacerbation was defined as a hospital admission for intravenous antibiotics. The investigators demonstrated that more than 50% of the decline in FEV1 was associated with one or more pulmonary exacerbation (fig. 2) [26]. Also, for a given number of exacerbations, the annual rate of FEV1 decline was greatest in subjects with $\leqslant$6 months between exacerbations [26]. Therefore, these findings would support using time to next exacerbation as a clinical end-point for clinical trials in CF.

However, there are some limitations to the use of pulmonary exacerbations as an end-point in clinical trials. Surprisingly, despite the significant clinical relevance of pulmonary exacerbations to patients with CF, at present, there is no agreed definition [27]. To date, various clinical trials that have used pulmonary exacerbations as an end-point have all adopted different definitions. Pulmonary exacerbations that have been linked to clinical outcomes have been the more severe exacerbations and the impact of less severe exacerbations is less well defined. Modified definitions have been used in patients with milder disease but these still need to be validated to assess the effect on FEV1 and survival. The threshold for using antibiotic therapy has become quite low and it has been recently demonstrated that physician-initiated antibiotic therapy not meeting the protocol-defined definition can affect the pulmonary exacerbation end-point in a clinical trial [28, 29].

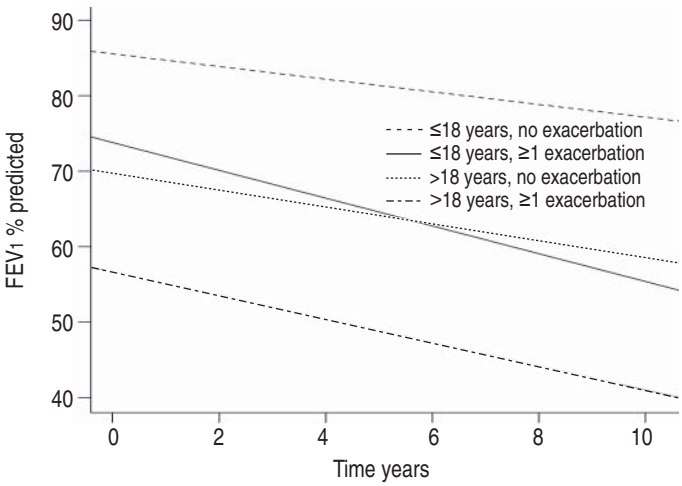


Figure 2. Forced expiratory volume in 1 s (FEV1) decline over time in subjects aged either $\leqslant$18 or >18 years, with or without exacerbation. Reproduced from [26].

Quality of life

Quality of life (QoL), as defined by the World Health Organization, is "the individual's perception of their position in life in the context of the culture and value systems in which they live in relation to their goals, expectations, standards and concerns" [30]. It has been previously shown in CF, as well as other chronic diseases, that QoL cannot be explained by clinical variables alone, thus adding credence to including PROs such as QoL measures in clinical trials [31].

Measures of QoL in CF patients have been linked to pulmonary function and, more recently, there is some evidence that it may also be linked to survival [31, 32]. The two most commonly used CF-specific QoL questionnaires are the CF Questionnaire Revised (CFQ-R) and the CF Quality of Life Questionnaire; both are validated, and have been shown to be reliable and sensitive [33, 34]. The CFQ-R has demonstrated treatment effects in large interventional trials [35–37]. The minimum clinically important difference (MCID) corresponds to the smallest clinically relevant change a patient can detect. The MCID has been determined for the respiratory domain of the CFQ-R using data from two clinical interventional trials for a population of stable CF patients, and was 4.0 and 8.5 for patients undergoing pulmonary exacerbations [38]. Notably, the MCID for the other domains and also for the overall CFQ-R has not yet been established.

The short-term efficacy and safety study of aztreonam lysine for inhalation in patients with CF with *P. aeruginosa* culture positivity was the first interventional trial that used a QoL measure, the CFQ-R, as the primary outcome. This was a randomised, double-blind, placebo-controlled, international study in patients $\geqslant$6 years of age with FEV_1 $\geqslant$25% and $\leqslant$75% pred [35]. Patients were treated with aztreonam (75 mg three times a day for 28 days) or placebo, and monitored for 14 days after study drug completion [35]. The investigators reported a significant improvement in the mean CFQ-R respiratory domain score (9.7 points ($p<0.001$) compared to placebo; fig. 3) [35]. However, it should be noted that this domain is largely a symptom score and not an overall measure of QoL. The CFQ-R respiratory domain was one of the secondary end-points in the ivacaftor study in CF patients >12 years of age as well as in the 6–11 years of age study [39, 40]. Interestingly, the CFQ-R respiratory domain significantly improved in the >12 years of age study but not in the study with the younger children, despite dramatic improvements in FEV_1 in both studies. This apparent difference between studies may reflect the differences in baseline respiratory severity in the two cohorts or the differences in sample size (n=161 *versus* n=52) [39, 40]. While the treatment effect in the ivacaftor 6–11 years of age study was greater than the MCID for the respiratory domain, it did not reach statistical significance due to its high variability in these patients with mild disease. Therefore, similar to the aforementioned limitations of FEV_1 and pulmonary exacerbations, the development and standardisation of new QoL instruments that are sensitive to smaller and earlier changes in symptoms are needed, given the overall improvement in lung function in consecutive CF birth cohorts [41].

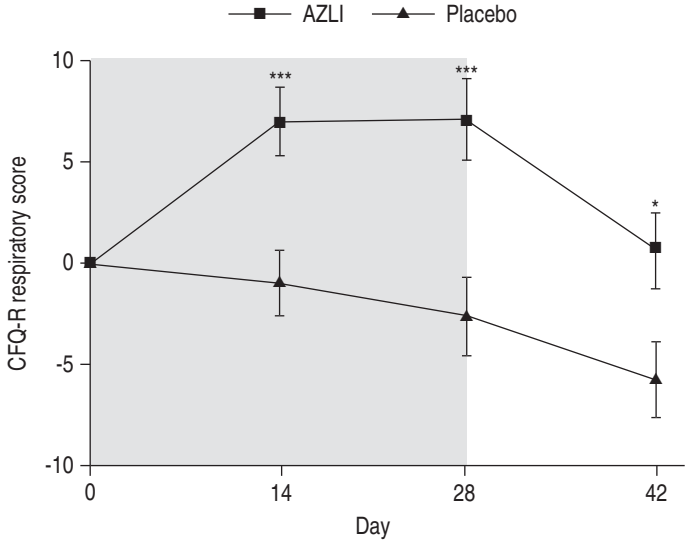


Figure 3. The adjusted mean Cystic Fibrosis Questionnaire Revised (CFQ-R) respiratory domain scores as change from baseline to study end (days 0–42). Child, teen and adult responses were combined for CFQ-R respiratory scores. Data are presented as mean ± SE. The shaded area represents the treatment period and the unshaded area represents the follow-up period. AZLI: inhaled aztreonam lysine. *: $p<0.05$ for AZLI *versus* placebo; ***: $p<0.001$ for AZLI *versus* placebo. Reproduced from [35] with permission from the publisher.

Novel end-points

Multiple-breath inert gas washout tests

Multiple-breath inert gas washout (MBW) tests assess the extent of ventilation inhomogeneity within the lungs. MBW tests were first described more than 50 years ago [42]. These tests involve the washout of an inert tracer gas from the lungs during relaxed, tidal breathing. A number of different inert gases can be used; resident gases, such as nitrogen, can be washed

out by breathing 100% oxygen, or nonresident gases, such as sulfur hexafluoride or helium, require an additional wash-in phase, after which they are washed out by breathing room air [43]. Airway narrowing due to mucus retention, airway inflammation and structural changes such as airway thickening can lead to ventilation inhomogeneity. This translates into a more prolonged washout, as a larger number of breaths are required for the washout of the tracer gas in the disease state. The indices of ventilation distribution efficiency from MBW can be divided into two categories: those describing global abnormalities (*e.g.* lung clearance index (LCI)) and those providing more detailed insight into the relative contributions of the main gas transport mechanisms within the lung, convection and diffusion [43]. This report will focus on global abnormalities, specifically LCI, which at present seems to be the most robust and sensitive, and hence the most reported in the CF literature.

One major advantage of MBW tests is their ability to be performed during tidal breathing and, therefore, to be performed by any age group. In addition, unlike FEV_1, the upper limits of normal for LCI are consistent across age except for the first years of life; published reference ranges are available for healthy control patients of all ages [17, 44–47]. The variability of LCI has also been studied across a wide age range and the coefficient of variation ranges from 7.8% for preschool children with CF to 3.2% for healthy adults [45, 47].

As mentioned previously, FEV_1 is an established clinical surrogate of mortality for CF patients [7]. Although, LCI has yet to be linked to mortality in longitudinal studies, it has been shown to be closely linked to FEV_1 in its ability to assess airways disease. The LCI appears to be sensitive to mild disease in the peripheral airways, the originating site of CF lung disease [2, 17, 44–49]. The LCI has repeatedly been shown to detect lung disease at an earlier stage than spirometry across a wide range of ages (fig. 4) [17, 44–47]. Further evidence of the LCI's validity comes from studies comparing LCI and spirometry with high-resolution CT (HRCT) scanning, thus bridging structural and physiological measures [50, 51]. In these studies, which included children at least 5 or 6 years of age, a normal LCI almost completely excluded CT abnormalities, thereby suggesting that serial LCI measurements could be used for early identification of structural lung disease [50, 51]. Similar studies have been performed in infants and younger children but the relationship between LCI and CT findings does not seem to be as clear [52, 53]. HALL *et al.* [53] found a weak but significant association between LCI and CT findings. OWENS *et al.* [52] reported that of nine children with normal LCI, five had abnormalities on the CT, whereas five children with normal HRCT had an elevated LCI. These results suggest that while the LCI and HRCT may have similar sensitivities in younger children with CF, complimentary information may be gained from both tests [53]. Future studies will need to address the longitudinal relationship between LCI and CT findings.

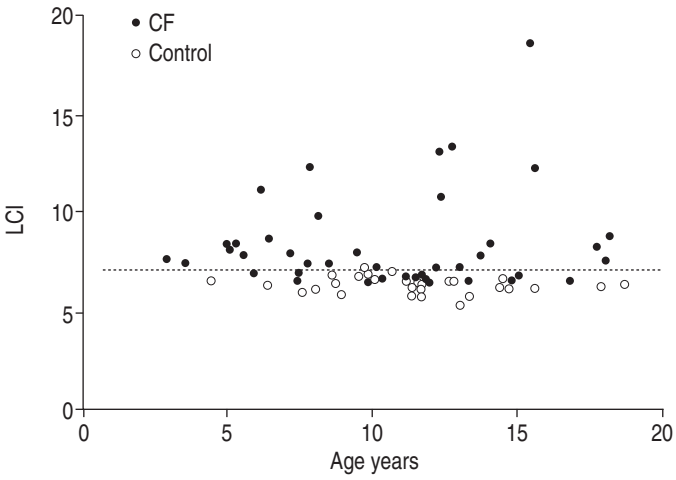


Figure 4. The lung clearance index (LCI) measured by the multiple-breath washout technique *versus* age in 43 patients with cystic fibrosis (CF) and 28 healthy controls. The dashed horizontal line represents the upper limit of normal for LCI. Reproduced from [17].

LCI also appears to be able to track disease progression over time [43, 48, 54]. A Swiss cohort of 142 CF children between the ages of 6 and 20 years were followed, and all had a minimum of four spirometric and LCI measurements per year [48]. LCI was the first lung function parameter to become abnormal and was more sensitive than all spirometric measures to disease progression [48]. The longitudinal

course of LCI from infancy to school age has also been demonstrated [54]. LCI measured during the preschool age range has been shown to be predictive for the development of abnormal spirometry at school age [55].

LCI has also been used as an outcome measure in interventional studies. ROBINSON *et al.* [56] demonstrated a significant improvement in LCI from intravenous antibiotic treatment for a pulmonary exacerbation on the LCI in CF patients with mild-to-moderate lung disease, although significant heterogeneity in the response was seen. Other interventional studies have demonstrated significant improvements in LCI but not FEV_1 after treatment with inhaled hypertonic saline and dornase alfa in two separate crossover design studies in which patients received 4 weeks of treatment or placebo in a randomised sequence separated by a 4-week washout period [57, 58]. Significant changes in LCI were seen despite small patient numbers, 20 patients [57] and 17 patients [58] in each of the studies. Early intervention studies such as the Infant Study of Inhaled Saline in Cystic Fibrosis (ISIS) required over 300 infants aged 4–60 months to power the primary end-point of pulmonary exacerbations [59]. A difference in pulmonary exacerbations was not seen between treatment arms in ISIS [59]. In contrast, a significant treatment effect was seen for LCI in the small subset of 25 patients that performed pre- and post-saline LCI [28]. LCI has also been shown to significantly improve in a cohort of patients bearing the G551D *CFTR* mutation with FEV_1 $\geqslant$90% pred for VX-770-106 [29]. Therefore, LCI seems to be particularly promising: significant treatment effects can be demonstrated in short periods and with small sample sizes.

However, there are some technical considerations that need to be addressed when discussing the use of the LCI as a primary outcome measure in multicentre clinical trials testing novel CF therapeutics. The majority of the aforementioned studies were performed using mass spectrometry and sulfur hexafluoride as a tracer gas, both of which are available in only a few centres. There is agreement within the CF community that nitrogen-based systems are more likely to be suitable for multicentre studies [43]. The Exhalyzer D system (ECO Medics AG, Dürnten, Switzerland) using nitrogen has been validated from late preschool to adult age; data are limited in infants and young children [60, 61]. The other commercially available system, EasyOne Pro (ndd, Zurich, Switzerland), has been validated over a range of volumes ranging from 350 to 4000 mL suggesting suitability from 2 years of age and up [43]. Therefore, the utility of both of these devices still needs to be clarified for younger children. Furthermore, the magnitude of effect of the inhalation of 100% oxygen on the breathing pattern of infants and MBW indices remains an area of needed future study. Additionally, more experience using nitrogen systems for preschool children is needed to determine feasibility and reproducibility as well as to obtain reference data [43]. Additionally, before LCI is "ready" for use as an outcome measure in multicentre trials, standard operating procedures need to be applied and adhered to across all centres. In summary, LCI is a very promising outcome for clinical trials in CF children with mild disease given its feasibility, sensitivity, reproducibility, responsiveness and ability to track change over time. However, further validation and experience in infants and preschool children with the commercially available nitrogen systems is still needed.

Chest CT

The use of chest CT as an outcome represents a paradigm shift towards structural disease and morphology as compared to the conventional focus on lung function as measured by FEV_1. CT scans appear to be a sensitive test for delineating early lung disease. In cohort studies of infants with CF, significant structural abnormalities, such as bronchiectasis, airway wall thickening and air trapping, were observed [62–64]. Additionally, despite the absence of respiratory symptoms in 84% (n=48) of infants in one of the cohorts, 80.7% (n=46) had an abnormal CT [64]. CT has also been shown to be more sensitive to early changes in CF lung disease, often preceding any change in lung function [15, 65, 66].

At present, chest CT is considered the gold standard for the detection of bronchiectasis [67]. Bronchiectasis is an important outcome parameter in CF because it is associated with a lower HRQoL and the presence of bronchiectasis is associated with a higher mortality [68]. Bronchiectasis

is the most validated finding in CT images from CF patients and seems to be the most reliable finding in young children [69]. Although air trapping is common early in life with CF, it is not known if this represents or predicts the onset of irreversible structural lung disease.

The responsiveness of chest CT to interventional studies has also been promising. Four studies have demonstrated significant improvements in CT scores after treatment for a pulmonary exacerbation in young children with CF as well as in older children and adult cohorts [70–73]. CT scores also improved after treatment with dornase alfa and tobramycin solution by inhalation in small pilot studies in paediatric patients with CF [74–77].

CT scans are able to track disease progression in longitudinal cohort studies [15, 16, 65]. Several studies have also evaluated the correlation of changes in CT with lung function parameters [15, 16, 66]. These seem to be weakly correlated, with a more progressive decline in CT scores as compared to pulmonary function [15, 16, 66]. The severity of structural changes on CT has also been shown to correlate with markers of inflammation and infection in BAL fluid [62, 64, 69, 73].

For manual scoring of the CT scans, a range of different systems are available with good reproducibility. In a comparative study, within- and between-reader reproducibility for five manual scoring systems was found to be good, with intraclass correlation coefficients $\geqslant 0.74$ [78]. However, these scoring systems require specific training and are highly time intensive. Fully automated scoring systems would be ideal as they would not require user input and would have near perfect intratest reproducibility, which would enable more rapid analyses. Unfortunately, validation in CF is lacking, with only one published study on automated scoring to date [79]. Therefore, further development is needed before automated systems can be used in clinical trials. Semiautomated systems fall in between the two previously mentioned approaches and have been validated in CF. These systems require the user to input the trachea, airway wall and vessel or thresholds for trapped air [80–83]. Semiautomated analyses are able to differentiate CF patients from healthy controls [81–83]. Additionally, quantitative trapped air measurements using semiautomated systems have been able to discriminate differences in treatment effect from dornase alfa *versus* placebo in a cohort of children with mild CF [74].

Radiation exposure and the cumulative dose for the paediatric population must also be considered in the discussion surrounding CT scans as endpoints for clinical trials in CF. Few studies have estimated accurately the radiation burden in patients with CF. De Gonzalez *et al.* [84] estimated cancer risk in patients with CF on the basis of conventional CT (low-dose studies of the chest performed annually starting at age 2 years). They found that lifelong annual CT examinations confer a low risk of radiation-induced cancer in CF; however, as the median overall CF survival estimate increases to 50 years of age, the risk of radiation-induced mortality modestly increases, with a lifetime risk of 0.08% for males and 0.46% for females [84]. Radiation doses continue to decrease with the development of newer CT scanners and improved analysis algorithms. However, these data indicate that radiation dose must be considered in imaging strategies for patients with CF to ensure that the benefits outweigh the risks [85–87].

To date, there have been no completed clinical trials that have used CT scans as primary end-point in CF. However, the ongoing COMBAT CF study (a randomised clinical trial of azithromycin to prevent bronchiectasis in infants; ClinicalTrials.gov identifier NCT01270074) is the first to use chest CT as a primary end-point [43]. Therefore, this study is a prototype and, if successful, it could be used as a template for future studies aimed at preventing structural lung disease [43].

FDG positron emission tomography/CT imaging

Fluorine-18-labelled 2-fluoro-2-deoxy-D-glucose (FDG) positron emission tomography (PET)/CT, a type of metabolic imaging, is another promising outcome measure. The uptake of FDG is highly correlated with cellular glucose metabolism by a process of metabolic trapping [88–90]. It has been shown to correlate with inflammation and the burden of neutrophils, the major cell type involved in CF lung disease [89, 91–93]. PET scanning can be used to determine the location and

intensity of FDG uptake in the lungs. Moreover, PET can be combined with CT to assess anatomy and structure with optimal resolution. Current technology houses a single combined PET/CT scanner to allow direct CT attenuation correction, as well as anatomic localisation of PET images [89, 94–98].

The first two PET imaging studies in patients with CF, those by LABIRIS *et al.* [88] and CHEN *et al.* [93] had mixed results and raised concerns about the sensitivity of PET in the detection of neutrophilic airway inflammation in CF. These investigators used tracer kinetic modelling, a quantitative form of analysis that looks at mean tracer uptake over a large lung area [99, 100]. More recent studies have applied the conventional semiquantitative method of obtaining standardised uptake values (SUVs), which is the most widely accepted method for quantifying FDG uptake [101]. This method supplies a static image, or "snapshot", of glucose metabolism, which is measured at a single point that takes into account several patient factors, such as patient weight and activity injected [96, 101, 102]. Furthermore, the use of the PET/CT technique has enabled correlation between PET and CT findings and the exclusion of mediastinal and hilar regions of lymph node activity [94–96, 101]. KLEIN *et al.* [103] reported the FDG PET/CT findings in a cross-sectional cohort of 20 adolescent and adult patients with CF. This group also examined FDG PET/CT studies before and after therapy for an exacerbation but small numbers and the different time points made it difficult to make any conclusions about the responsiveness of the tool. More recently, AMIN *et al.* [104] conducted FDG PET/CT examinations (with low-dose CT) in 20 children with CF before and after 14 days of intravenous antibiotic treatment for a pulmonary exacerbation. PET activity was quantified by using SUVs through assessment of background activity (SUV_{mean}) and superimposed focal uptake (SUV_{max}) for each lung zone. Antibiotic therapy resulted in a significant decrease in SUV_{max}. CT scores significantly correlated with SUV_{max}. Sputum neutrophils correlated with SUV_{mean}. The investigators demonstrated that FDG PET/CT is capable of detecting changes in inflammation in the lung after intravenous antibiotic treatment for a pulmonary exacerbation. These results show that different values generated from FDG PET/CT provide information on different aspects of CF (*i.e.* diffuse and focal inflammation). The pre-treatment studies demonstrated an increased background lung activity with superimposed multifocal activity. After therapy, SUV_{mean} normalised, and SUV_{max} improved but did not fully resolve (fig. 5). This partial improvement, however, correlated positively with overall lung function improvement. As it is well established that airway inflammation persists over time in patients with CF, SUV_{max} may be a more useful outcome measure for anti-inflammatory interventions compared with the overall activity. Therefore, although FDG PET/CT imaging can help quantify and localise pulmonary inflammation in CF, and track clinical improvement in patients with CF, it has not yet been fully evaluated in terms of radiation exposure risks *versus* actual benefits.

Pulmonary magnetic resonance imaging

Given the concerns with radiation exposure in both CT imaging and FDG PET/CT, magnetic resonance imaging (MRI) has recently been established as a radiation-free alternative. MRI of the chest was first proposed as a potential imaging modality in patients with CF in the late 1980s, although at that time, MRI was not capable of the same resolution of the pulmonary parenchyma as CT [105, 106]. There have been significant technological advances in MRI leading to improved temporal and/or spatial resolution [107–109]. There is some evidence that MRI is comparable to CT with regard to the detection of larger morphological changes in the CF lung [110–112]. However, the spatial resolution of MRI remains lower than CT and normal peripheral bronchi at the third to fourth generation are poorly visualised [113].

One of the main advantages of MRI over CT is the additional assessment of "function", *i.e.* perfusion, pulmonary haemodynamics and ventilation [113]. MRI has the advantage of being able to capture different regional functional aspects of the lung parenchyma as well as different aspects of the tissue based on different contrasts. In CF, regional ventilatory defects result in changes in regional lung perfusion due to the reflex of hypoxic vasoconstriction or tissue destruction [113]. MRI is able to assess lung perfusion using techniques with or without contrast agents [114–116]. In CF children, perfusion defects have been shown to correlate with tissue destruction; changes in

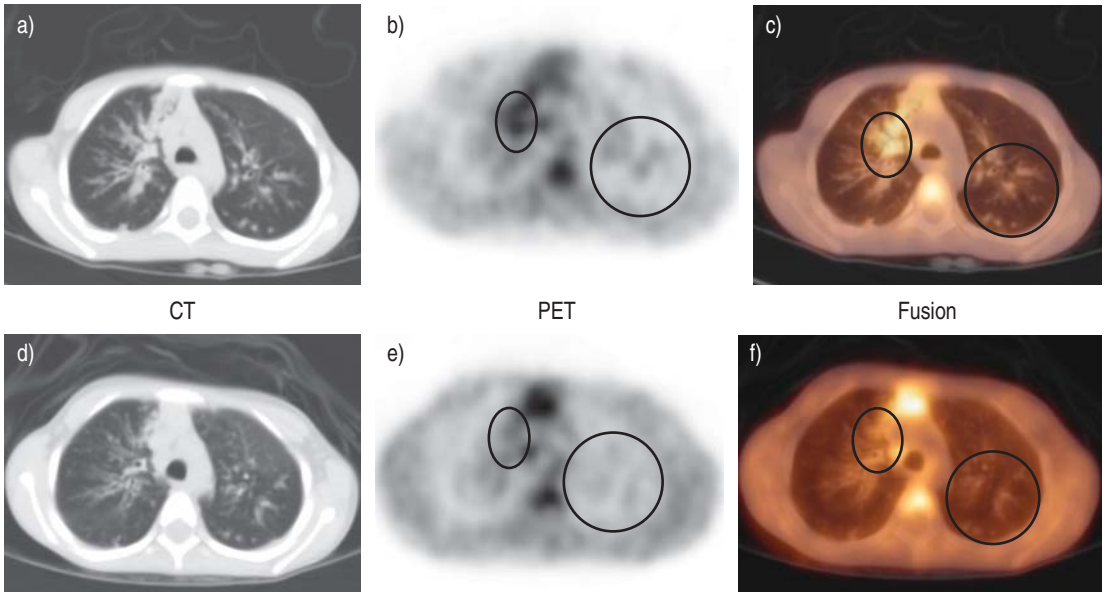


Figure 5. a–c) Pre- and d–f) post-therapy, a, d) computed tomography (CT), b, e) positron emission tomography (PET) and c, f) PET/CT fusion images in a study subject with cystic fibrosis. Ovals: maximal standardised uptake value (SUV); circles: mean SUV. Reproduced from [104] with permission from the publisher.

lung perfusion in CF children aged 0–6 years have also been observed in the absence of morphological changes [113]. A promising technique for imaging ventilation and perfusion of the lung based on Fourier decomposition, not relying on the injection of an intravenous contrast agent, was recently described [117]. This technique is particularly promising for longitudinal studies because it does not depend on any breath holding or triggering/gating technique [118]. MRI also allows for ventilation imaging using the inhalation of different gases (oxygen, helium-3 and xenon-129) for contrast enhancement. In a study using helium-3, patients with CF had a significantly higher number of ventilation defects than healthy controls [119]. The study also demonstrated the feasibility of using hyperpolarised helium-3 as an outcome measure in the evaluation of airway clearance techniques [119].

However, at present, standardised and validated operating procedures do not exist for MRI in CF patients. Although a lot of work still lies ahead before MRI imaging can be used as an end-point for clinical trials in CF, it is promising given the absence of radiation exposure, ability to detect functional as well as morphological changes and feasibility to be used by children of all ages.

Bronchoalveolar lavage

CF is characterised by recurrent cycles of infection and inflammation that begin early in life [120]. At present, BAL is the gold standard for detecting lower airway inflammation and infection in younger children with CF [121]. The accuracy of defining patients' infectious states, especially in relation to early infection with *Staphylococcus aureus*, has been shown to be superior to other available culture techniques in non-sputum producing infants and children [122]. BAL studies in infants with CF have identified elevated neutrophil counts in inflamed CF airways compared with those in control subjects [10, 21, 123, 124]. The Australian Early Surveillance Team for Cystic Fibrosis (AREST CF) has shown that infection and inflammation are predictive of subsequent airway disease quantified by CT [69]. Furthermore, these investigators have recently demonstrated that free neutrophil elastase from BAL samples in infants with CF is associated with persistent bronchiectasis (present on two or more sequential scans), with the odds seven times as high at 12 months of age and four times as high at 3 years of age [125].

There are some limitations when considering BAL as an outcome measure in clinical trials in CF. CF lung disease is heterogeneous. Therefore, BAL performed in one lobe of the lung may not be representative of the pathology in the rest of the lobes. Standard operating procedures for a BAL have been proposed by a European Respiratory Society Task Force and are in the process of being updated [121]. However, despite this, the ideal BAL protocol with respect to obtaining and processing samples has not yet been defined. Furthermore, in the majority of paediatric centres worldwide, BAL requires a general anaesthetic. Therefore, its invasiveness precludes its use as a primary outcome measure in clinical trials in infants and young children. However, BAL plays an important role in cross-validating other less invasive biomarkers, such as sputum, that could potentially be used as outcome measures that could be more frequently sampled in an interventional trial in CF.

Sputum

Biomarkers of inflammation and infection in CF patients are noninvasive measures that can be quantified, in sputum samples. However, validation is paramount before biomarkers should be used as primary end-points in clinical trials. Just over 50% of children aged 6–18 years are able to expectorate sputum [126]. An additional 20% of children aged 6–18 years are able to provide a sputum sample after sputum induction [127]. Some centres have also had success with sputum induction in preschool children [128, 129]. However, the majority of children under 5 years of age are not able to produce either expectorated or induced sputum. This is a potential limitation of the use of sputum biomarkers as end-points in clinical trials. One study in stable, adult CF patients has shown that induced sputum samples appeared to be equally sensitive compared with expectorated sputum or BAL samples in detecting bacterial pathogens [130]. Bronchospasm is the main adverse effect of sputum induction and seems to occur in about 7% of patients [131–134].

Support for sputum samples as biomarkers of disease first emerged from small-cohort, single-centre studies. SAGEL *et al.* [135] reported a small cohort of 19 CF patients with mild lung disease and showed that pulmonary function and induced sputum measurements of interleukin (IL)-8, neutrophil elastase, total cell counts and neutrophil counts were significantly correlated. KIM *et al.* [136] also demonstrated a significant correlation between pulmonary function and IL-8 and myeloperoxidase in 16 patients with CF. MAYER-HAMBLETT *et al.* [137] built upon these results and pooled the sputum data from four studies performed using the Cystic Fibrosis Foundation Therapeutic Drug Network (CFF TDN). The aim was to quantify the cross-sectional association between FEV_1 and expectorated sputum markers to determine which sputum biomarkers can best explain the interpatient variation in pulmonary function. Overall lung function correlated negatively with sputum free elastase, IL-8, neutrophil counts and percentage of neutrophils. Free elastase had the greatest correlation with pulmonary function (-0.35, 95% CI -0.46– -0.22) and explained 12% of the interpatient variability in lung function. Elastase was further shown to have a significant longitudinal association with FEV_1, specifically a -2.9% decline in FEV_1 (95% CI -5.0– -0.9) per 1-log increase in elastase. SAGEL *et al.* [138] examined the longitudinal relationships between sputum biomarkers and lung function. The investigators performed sputum induction annually for 3 years in 35 children with CF; sputum was assayed for mediators related to proteolysis and a panel of inflammatory cytokines. Higher detectable baseline neutrophil elastase was associated with more rapid lung function decline. Furthermore, when categorising subjects as "rapid" or "slow" decliners, logistic regression, demonstrated that the initial measurement of neutrophil elastase had the highest individual predictive value for subsequent lung function, whereas neutrophil elastase, IL-8 and IL-6 had the highest combined predictive value. Therefore, in children with CF, a single determination of sputum biomarkers, particularly neutrophil elastase, has predictive value for subsequent lung function decline. This was the first study to demonstrate that longitudinal changes in sputum inflammatory biomarkers are related to lung function changes. Other sputum biomarkers that are currently being explored include messenger RNA, HMB-1 and the α_1-antitrypsin:CD16b complex [139, 140]. Therefore, sputum biomarkers are a promising clinical end-point for interventional studies of anti-inflammatory therapies but are limited by the age of the patients in which they can be performed.

Intestinal current measurement

ICM measures the electric current resulting from epithelial ion fluxes at the mucosal surface, *ex vivo* [141]. The intestinal sample is usually a rectal suction biopsy and a special micro-Ussing chamber is needed to measure the transepithelial short-circuit current (I_{sc}) [141]. In CF, the intestinal CFTR-mediated chloride secretion is impaired, while absorptive processes remain unchanged or may be enhanced [141]. The normal I_{sc} response to forskolin, an activator of CFTR, is absent or reduced in CF. In ICMs of healthy individuals, the apical potassium efflux in response to carbachol and histamine is masked by the much larger chloride efflux [141]. In CF, the response is reversed due to the apical potassium efflux in the absence of a chloride efflux or biphasic due to residual CFTR-mediated chloride efflux in milder forms of CF [142–144].

The validity of the ICM has been previously demonstrated. ICM has been shown to discriminate between patients with CF and healthy individuals [142, 145–151] and, at a group level, can discriminate between pancreatic-sufficient and -insufficient patients [143]. ICM is particularly promising as a clinical end-point because it is the only biomarker at present that is able to directly assess the beneficial effects of CFTR correctors in patient-derived tissue *ex vivo* [152]. Therefore, strong potentiators can be applied directly to the tissue removed from the patient under corrector treatment to assess full CFTR activity [141]. Evidence for discriminant validity comes from the finding that more chloride secretion measured in the rectal mucosa correlates with a milder disease presentation [143, 144, 147]. Additionally, ICM has been shown to strongly correlate with *CFTR* mutation analysis and moderately with sweat chloride [142]. However, further data are needed on the reliability and responsiveness of ICM. Furthermore, to date, no studies have used ICM as an end-point in clinical trials.

At present, ICMs are limited to a few centres with the appropriate expertise, given the complexity of the measurements. Therefore, the development of standard operating procedures with strict adherence to the protocol is needed if ICMs are to successfully transition from biomarker to clinical end-point.

Conclusion

In summary, a wide range of promising biomarkers are available that measure different aspects of CF lung disease ranging from morphological markers to those measuring lung function and inflammation. It is unlikely that one method alone will become the "holy grail" technique to assess lung disease but rather that different outcomes will be needed based on the intervention being studied [43]. MBW and CT imaging seem to be the furthest along in the validation process but work still needs to be done for these end-points, particularly regarding standardisation. With the paradigm shift towards interventions targeting younger children with milder CF, ensuring adequate sensitivity, validation and clinical relevance of up-and-coming biomarkers is essential for future clinical trials and highly needed to further advance CF care.

Statement of Interest

F. Ratjen has acted as a consultant for Vertex on CF-related activities. He currently acts as a consultant for Novartis, Bayer, Talecris, CSL Behring, Roche and Gilead on CF-related activities. He is the principal investigator on a pending grant from Novartis. He has developed a speaker's bureau for Genentech Inc. and has given talks within this programme. He has given talks for and received travel expenses from Pari.

References

1. Decramer M, Rutten-van Mölken M, Dekhuijzen PN, *et al.* Effects of N-acetylcysteine on outcomes in chronic obstructive pulmonary disease (Bronchitis Randomized on NAC Cost-Utility Study, BRONCUS): a randomised placebo-controlled trial. *Lancet* 2005; 365: 1552–1560.
2. Ratjen F, Doring G. Cystic fibrosis. *Lancet* 2003; 361: 681–689.

3. Gibson RL, Burns JL, Ramsey BW. Pathophysiology and management of pulmonary infections in cystic fibrosis. *Am J Respir Crit Care Med* 2003; 168: 918–951.
4. Chmiel JF, Berger M, Konstan MW. The role of inflammation in the pathophysiology of CF lung disease. *Clin Rev Allergy Immunol* 2002; 23: 5–27.
5. Robinson MB. Mucociliary clearance in cystic fibrosis. *Pediatr Pulmonol* 2002; 33: 293–306.
6. Katz R. Biomarkers and surrogate markers: an FDA perspective. *NeuroRx* 2004; 1: 189–195.
7. Corey M, Farewell V. Determinants of mortality from cystic fibrosis in Canada, 1970–1989. *Am J Epidemiol* 1996; 143: 1007–1017.
8. Corey M, Edwards L, Levison H, *et al.* Longitudinal analysis of pulmonary function decline in patients with cystic fibrosis. *J Pediatr* 1997; 131: 809–814.
9. Konstan MW, Hilliard KA, Norvell TM, *et al.* Bronchoalveolar lavage findings in cystic fibrosis patients with stable, clinically mild lung disease suggest ongoing infection and inflammation. *Am J Respir Crit Care Med* 1994; 150: 448–454.
10. Khan TZ, Wagener JS, Bost T, *et al.* Early pulmonary inflammation in infants with cystic fibrosis. *Am J Respir Crit Care Med* 1995; 151: 1075–1082.
11. Dakin CJ, Numa AH, Wang H, *et al.* Inflammation, infection, and pulmonary function in infants and young children with cystic fibrosis. *Am J Respir Crit Care Med* 2002; 165: 904–910.
12. Ranganathan SC, Dezateux C, Bush A, *et al.* Airway function in infants newly diagnosed with cystic fibrosis. *Lancet* 2001; 358: 1964–1965.
13. Accurso FJ. Introduction: the rational for early intervention in cystic fibrosis. *Pediatr Pulmonol* 1997; 24: 132–133.
14. Gelb AF, Williams AJ, Zamel N. Spirometry. FEV_1 *vs* $FEF_{25–75}$ percent. *Chest* 1983; 84: 473–474.
15. de Jong PA, Nakano Y, Leguin MH, *et al.* Progressive damage on high resolution computed tomography despite stable lung function in cystic fibrosis. *Eur Respir J* 2004; 23: 93–97.
16. de Jong PA, Lindblad A, Rubin L, *et al.* Progression of lung disease on computed tomography and pulmonary function tests in children and adults with cystic fibrosis. *Thorax* 2006; 61: 80–85.
17. Gustafsson PM, Aurora P, Lindblad A. Evaluation of ventilation maldistribution as an early indicator of lung disease in children with cystic fibrosis. *Eur Respir J* 2003; 22: 972–979.
18. Aurora P, Wade A, Whitmore P, *et al.* A model for predicting life expectancy of children with cystic fibrosis. *Eur Respir J* 2000; 16: 1056–1060.
19. Macklem PT, Mead J. Resistance of central and peripheral airways measured by a retrograde catheter. *J Appl Physiol* 1967; 22: 395–401.
20. Hogg JC, Macklem PT, Thurlbeck WM. Site and nature of airway obstruction in chronic obstructive lung disease. *N Engl J Med* 1968; 278: 1355–1360.
21. Goss CH, Burns JL. Exacerbations in cystic fibrosis. 1: Epidemiology and pathogenesis. *Thorax* 2007; 62: 360–367.
22. Flume PA, Mogayzel PJ Jr, Robinson KA, *et al.* Cystic fibrosis pulmonary guidelines: treatment of pulmonary exacerbations. *Am J Respir Crit Care Med* 2009; 180: 802–808.
23. Mayer-Hamblett N, Rosenfeld M, Emerson J, *et al.* Developing cystic fibrosis lung transplant referral criteria using predictors of 2-year mortality. *Am J Respir Crit Care Med* 2002; 166: 1550–1555.
24. Liou TG, Adler FR, Fitzsimmons SC, *et al.* Predictive 5-year survivorship model of cystic fibrosis. *Am J Epidemiol* 2001; 153: 345–352.
25. Konstan MW, Morgan WJ, Butler SM, *et al.* Risk factors for rate of decline in forced expiratory volume in one second in children and adolescents with cystic fibrosis. *J Pediatr* 2007; 151: 134–139.
26. Waters V, Stanojevic S, Atenafu EG, *et al.* Effect of pulmonary exacerbations on long-term lung function decline in cystic fibrosis. *Eur Respir J* 2012; 40: 61–66.
27. Rosenfeld M, Emerson J, Williams-Warren J, *et al.* Defining a pulmonary exacerbation in cystic fibrosis. *J Pediatr* 2001; 139: 359–365.
28. Subbarao P, Stanojevic S, Brown M, *et al.* Lung clearance index as an outcome measure for clinical trials in young children with cystic fibrosis: a pilot study using inhaled hypertonic saline. *Am J Respir Crit Care Med* 2013; 188: 456–460.
29. Ratjen FSH, Lee PS, Song T, *et al.* Effect of ivacaftor on lung clearance index and FEV in subjects with CF who have the G551D CFTR utation and mild lung disease. *Eur Respir J* 2012; 40: Suppl. 56, 316s.
30. WHOQOL Group. Development of the WHOQOL: rationale and current status. *Int J Ment Health* 1994; 23: 24–56.
31. Gee L, Abbott J, Hart A, *et al.* Associations between clinical variables and quality of life in adults with cystic fibrosis. *J Cyst Fibros* 2005; 4: 59–66.
32. Abbott J. Health-related quality of life measurement in cystic fibrosis: advances and limitations. *Chron Respir Dis* 2009; 6: 31–41.
33. Quittner AL, Buu A, Messer MA, *et al.* Development and validation of The Cystic Fibrosis Questionnaire in the United States: a health-related quality-of-life measure for cystic fibrosis. *Chest* 2005; 128: 2347–2354.
34. Dobbin CJ, Bartlett D, Melehan K, *et al.* The effect of infective exacerbations on sleep and neurobehavioral function in cystic fibrosis. *Am J Respir Crit Care Med* 2005; 172: 99–104.
35. Retsch-Bogart GZ, Quittner AL, Gibson RL, *et al.* Efficacy and safety of inhaled aztreonam lysine for airway *Pseudomonas* in cystic fibrosis. *Chest* 2009; 135: 1223–1232.

36. Donaldson SH, Bennett WD, Zeman KL, *et al.* Mucus clearance and lung function in cystic fibrosis with hypertonic saline. *N Engl J Med* 2006; 354: 241–250.
37. Elkins MR, Robinson M, Rose BR, *et al.* A controlled trial of long-term inhaled hypertonic saline in patients with cystic fibrosis. *N Engl J Med* 2006; 354: 229–240.
38. Quittner AL, Modi AC, Wainwright C, *et al.* Determination of the minimal clinically important difference scores for the Cystic Fibrosis Questionnaire-Revised respiratory symptom scale in two populations of patients with cystic fibrosis and chronic *Pseudomonas aeruginosa* airway infection. *Chest* 2009; 135: 1610–1618.
39. Ramsey BW, Davies J, McElvaney NG, *et al.* A CFTR potentiator in patients with cystic fibrosis and the G551D mutation. *N Engl J Med* 2011; 365: 1663–1672.
40. Davies JC, Wainwright CE, Canny GJ, *et al.* Efficacy and safety of ivacaftor in patients aged 6 to 11 years with cystic fibrosis with a G551D mutation. *Am J Respir Crit Care Med* 2013; 187: 1219–1225.
41. Goss CH, Quittner AL. Patient-reported outcomes in cystic fibrosis. *Proc Am Thorac Soc* 2007; 4: 378–386.
42. Becklake MR. A new index of the intrapulmonary mixture of inspired air. *Thorax* 1952; 7: 111–116.
43. Stick S, Tiddens H, Aurora P, *et al.* Early intervention studies in infants and preschool children with cystic fibrosis: are we ready? *Eur Respir J* 2013; 42: 527–538.
44. Aurora P, Gustafsson P, Bush A, *et al.* Multiple breath inert gas washout as a measure of ventilation distribution in children with cystic fibrosis. *Thorax* 2004; 59: 1068–1073.
45. Aurora P, Bush A, Gustafsson P, *et al.* Multiple-breath washout as a marker of lung disease in preschool children with cystic fibrosis. *Am J Respir Crit Care Med* 2005; 171: 249–256.
46. Lum S, Gustafsson P, Ljungberg H, *et al.* Early detection of cystic fibrosis lung disease: multiple-breath washout *versus* raised volume tests. *Thorax* 2007; 62: 341–347.
47. Horsley AR, Gustafsson PM, Macleod KA, *et al.* Lung clearance index is a sensitive, repeatable and practical measure of airways disease in adults with cystic fibrosis. *Thorax* 2008; 63: 135–140.
48. Kraemer R, Blum A, Schibler A, *et al.* Ventilation inhomogeneities in relation to standard lung function in patients with cystic fibrosis. *Am J Respir Crit Care Med* 2005; 171: 371–378.
49. Davies JC, Cunningham S, Alton EW, *et al.* Lung clearance index in CF: a sensitive marker of lung disease severity. *Thorax* 2008; 63: 96–97.
50. Gustafsson PM, De Jong PA, Tiddens HA, *et al.* Multiple-breath inert gas washout and spirometry *versus* structural lung disease in cystic fibrosis. *Thorax* 2008; 63: 129–134.
51. Ellemunter H, Fuchs SI, Unsinn KM, *et al.* Sensitivity of lung clearance index and chest computed tomography in early CF lung disease. *Respir Med* 2010; 104: 1834–1842.
52. Owens CM, Aurora P, Stanojevic S, *et al.* Lung clearance index and HRCT are complementary markers of lung abnormalities in young children with CF. *Thorax* 2011; 66: 481–488.
53. Hall GL, Logie KM, Parsons F, *et al.* Air trapping on chest CT is associated with worse ventilation distribution in infants with cystic fibrosis diagnosed following newborn screening. *PLoS One* 2011; 6: e23932.
54. Kieninger E, Singer F, Fuchs O, *et al.* Long-term course of lung clearance index between infancy and school-age in cystic fibrosis subjects. *J Cyst Fibros* 2011; 10: 487–490.
55. Aurora P, Stanojevic S, Wade A, *et al.* Lung clearance index at 4 years predicts subsequent lung function in children with cystic fibrosis. *Am J Respir Crit Care Med* 2011; 183: 752–758.
56. Robinson PD, Cooper P, Van Asperen P, *et al.* Using index of ventilation to assess response to treatment for acute pulmonary exacerbation in children with cystic fibrosis. *Pediatr Pulmonol* 2009; 44: 733–742.
57. Amin R, Subbarao P, Jabar A, *et al.* Hypertonic saline improves the LCI in paediatric patients with CF with normal lung function. *Thorax* 2010; 65: 379–383.
58. Amin R, Subbarao P, Lou W, *et al.* The effect of dornase alfa on ventilation inhomogeneity in patients with cystic fibrosis. *Eur Respir J* 2011; 37: 806–812.
59. Rosenfeld M, Ratjen F, Brumback L, *et al.* Inhaled hypertonic saline in infants and children younger than 6 years with cystic fibrosis: the ISIS randomized controlled trial. *JAMA* 2012; 307: 2269–2277.
60. Singer F, Houltz B, Latzin P, *et al.* A realistic validation study of a new nitrogen multiple-breath washout system. *PLoS One* 2012; 7: e36083.
61. Houltz B, Green K, Lindblad A, *et al.* Tidal N_2 washout ventilation inhomogeneity indices in a reference population aged 7–70 years. *Eur Respir J* 2012; 40: Suppl. 56, 694s.
62. Stick SM, Brennan S, Murray C, *et al.* Bronchiectasis in infants and preschool children diagnosed with cystic fibrosis after newborn screening. *J Pediatr* 2009; 155: 623–628.
63. Martinez TM, Llapur CJ, Williams TH, *et al.* High-resolution computed tomography imaging of airway disease in infants with cystic fibrosis. *Am J Respir Crit Care Med* 2005; 172: 1133–1138.
64. Sly PD, Brennan S, Gangell C, *et al.* Lung disease at diagnosis in infants with cystic fibrosis detected by newborn screening. *Am J Respir Crit Care Med* 2009; 180: 146–152.
65. Helbich TH, Heinz-Peer G, Fleischmann D, *et al.* Evolution of CT findings in patients with cystic fibrosis. *AJR Am J Roentgenol* 1999; 173: 81–88.
66. Brody AS, Klein JS, Molina PL, *et al.* High-resolution computed tomography in young patients with cystic fibrosis: distribution of abnormalities and correlation with pulmonary function tests. *J Pediatr* 2004; 145: 32–38.
67. Hansell DM. Bronchiectasis. *Radiol Clin North Am* 1998; 36: 107–128.

68. Loeve M, Hop WC, de Bruijne M, *et al.* Chest computed tomography scores are predictive of survival in patients with cystic fibrosis awaiting lung transplantation. *Am J Respir Crit Care Med* 2012; 185: 1096–1103.
69. Mott LS, Park J, Murray CP, *et al.* Progression of early structural lung disease in young children with cystic fibrosis assessed using CT. *Thorax* 2012; 67: 509–516.
70. Shah RM, Sexauer W, Ostrum BJ, *et al.* High-resolution CT in the acute exacerbation of cystic fibrosis: evaluation of acute findings, reversibility of those findings, and clinical correlation. *AJR Am J Roentgenol* 1997; 169: 375–380.
71. Robinson TE, Leung AN, Northway WH, *et al.* Spirometer-triggered high-resolution computed tomography and pulmonary function measurements during an acute exacerbation in patients with cystic fibrosis. *J Pediatr* 2001; 138: 553–559.
72. Brody AS, Molina PL, Klein JS, *et al.* High-resolution computed tomography of the chest in children with cystic fibrosis: support for use as an outcome surrogate. *Pediatr Radiol* 1999; 29: 731–735.
73. Davis SD, Fordham LA, Brody AS, *et al.* Computed tomography reflects lower airway inflammation and tracks changes in early cystic fibrosis. *Am J Respir Crit Care Med* 2007; 175: 943–950.
74. Robinson TE, Goris ML, Zhu HJ, *et al.* Dornase alfa reduces air trapping in children with mild cystic fibrosis lung disease: a quantitative analysis. *Chest* 2005; 128: 2327–2335.
75. Robinson TE, Leung AN, Northway WH, *et al.* Composite spirometric-computed tomography outcome measure in early cystic fibrosis lung disease. *Am J Respir Crit Care Med* 2003; 168: 588–593.
76. Nasr SZ, Kuhns LR, Brown RW, *et al.* Use of computerized tomography and chest x-rays in evaluating efficacy of aerosolized recombinant human DNase in cystic fibrosis patients younger than age 5 years: a preliminary study. *Pediatr Pulmonol* 2001; 31: 377–382.
77. Nasr SZ, Gordon D, Sakmar E, *et al.* High resolution computerized tomography of the chest and pulmonary function testing in evaluating the effect of tobramycin solution for inhalation in cystic fibrosis patients. *Pediatr Pulmonol* 2006; 41: 1129–1137.
78. de Jong PA, Tiddens HA, Lequin MH, *et al.* Estimation of the radiation dose from CT in cystic fibrosis. *Chest* 2008; 133: 1289–1291.
79. Tschirren J, Hoffman EA, McLennan G, *et al.* Segmentation and quantitative analysis of intrathoracic airway trees from computed tomography images. *Proc Am Thorac Soc* 2005; 2: 484–487.
80. Mumcuoglu EU, Prescott J, Baker BN, *et al.* Image analysis for cystic fibrosis: automatic lung airway wall and vessel measurement on CT images. *Conf Proc IEEE Eng Med Biol Soc* 2009; 2009: 3545–3548.
81. Montaudon M, Berger P, Cangini-Sacher A, *et al.* Bronchial measurement with three-dimensional quantitative thin-section CT in patients with cystic fibrosis. *Radiology* 2007; 242: 573–581.
82. Goris ML, Zhu HJ, Blankenberg F, *et al.* An automated approach to quantitative air trapping measurements in mild cystic fibrosis. *Chest* 2003; 123: 1655–1663.
83. Bonnel AS, Song SM, Kesavarju K, *et al.* Quantitative air-trapping analysis in children with mild cystic fibrosis lung disease. *Pediatr Pulmonol* 2004; 38: 396–405.
84. de Gonzalez AB, Kim KP, Samet JM. Radiation-induced cancer risk from annual computed tomography for patients with cystic fibrosis. *Am J Respir Crit Care Med* 2007; 176: 970–973.
85. Tiddens HA. Chest computed tomography scans should be considered as a routine investigation in cystic fibrosis. *Paediatr Respir Rev* 2006; 7: 202–208.
86. Brenner DJ. Estimating cancer risks from pediatric CT: going from the qualitative to the quantitative. *Pediatr Radiol* 2002; 32: 228–223.
87. de Jong PA, Song SM, Kesavarju K, *et al.* Estimation of cancer mortality associated with repetitive computed tomography scanning. *Am J Respir Crit Care Med* 2006; 173: 199–203.
88. Labiris NR, Nahmias C, Freitag AP, *et al.* Uptake of ^{18}fluorodeoxyglucose in the cystic fibrosis lung: a measure of lung inflammation? *Eur Respir J* 2003; 21: 848–854.
89. Chen DL, Schuster DP. Imaging pulmonary inflammation with positron emission tomography: a biomarker for drug development. *Mol Pharm* 2006; 3: 488–495.
90. Zhuang H, Alavi A. 18-fluorodeoxyglucose positron emission tomographic imaging in the detection and monitoring of infection and inflammation. *Semin Nucl Med* 2002; 32: 47–59.
91. Chen DL, Rosenbluth DB, Mintun MA, *et al.* FDG-PET imaging of pulmonary inflammation in healthy volunteers after airway instillation of endotoxin. *J Appl Physiol* 2006; 100: 1602–1609.
92. Chacko TK, Zhuang H, Nakhoda KZ, *et al.* Applications of fluorodeoxyglucose positron emission tomography in the diagnosis of infection. *Nucl Med Commun* 2003; 24: 615–624.
93. Chen DL, Ferkol TW, Mintun MA, *et al.* Quantifying pulmonary inflammation in cystic fibrosis with positron emission tomography. *Am J Respir Crit Care Med* 2006; 173: 1363–1369.
94. Czernin J, Schelbert H. PET/CT imaging: facts, opinions, hopes, and questions. *J Nucl Med* 2004; 45: 1S–3S.
95. Kapoor V, McCook BM, Torok FS. An introduction to PET-CT imaging. *Radiographics* 2004; 24: 523–543.
96. Freeman LM, Blaufox MD. Letter from the Editors. *Semin Nucl Med* 2004; 34: 83.
97. Kostakoglu L, Hardoff R, Mirtcheva R, *et al.* PET-CT fusion imaging in differentiating physiologic from pathologic FDG uptake. *Radiographics* 2004; 24: 1411–1431.
98. Mettler FA Jr, Guiberteau MJ. Essentials of Nuclear Medicine Imaging. Philadelphia, Saunders, 2011.
99. Patlak CS, Blasberg RG. Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. Generalizations. *J Cereb Blood Flow Metab* 1985; 5: 584–590.

100. Patlak CS, Blasberg RG, Fenstermacher JD. Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. *J Cereb Blood Flow Metab* 1983; 3: 1–7.
101. Dolovich MB, Schuster DP. Positron emission tomography and computed tomography *versus* positron emission tomography–computed tomography. *Proc Am Thorac Soc* 2007; 4: 328–333.
102. Hunter GJ, Hamberg LM, Alpert NM, *et al.* Simplified measurement of deoxyglucose utilization rate. *J Nucl Med* 1996; 37: 950–955.
103. Klein M, Cohen-Cymberknoh M, Armoni S, *et al.* ^{18}F-fluorodeoxyglucose-PET/CT imaging of lungs in patients with cystic fibrosis. *Chest* 2009; 136: 1220–1228.
104. Amin R, Charron M, Grinblat L, *et al.* Cystic fibrosis: detecting changes in airway inflammation with FDG PET/CT. *Radiology* 2012; 264: 868–875.
105. Fiel SB, Friedman AC, Caroline DF, *et al.* Magnetic resonance imaging in young adults with cystic fibrosis. *Chest* 1987; 91: 181–184.
106. Carr DH, Oades P, Trotman-Dickenson B, *et al.* Magnetic resonance scanning in cystic fibrosis: comparison with computed tomography. *Clin Radiol* 1995; 50: 84–89.
107. Swift AJ, Woodhouse N, Fichele S, *et al.* Rapid lung volumetry using ultrafast dynamic magnetic resonance imaging during forced vital capacity maneuver: correlation with spirometry. *Invest Radiol* 2007; 42: 37–41.
108. Heidemann RM, Griswold MA, Kiefer B, *et al.* Resolution enhancement in lung ^{1}H imaging using parallel imaging methods. *Magn Reson Med* 2003; 49: 391–394.
109. Fink C, Bock M, Puderbach M, *et al.* Partially parallel three-dimensional magnetic resonance imaging for the assessment of lung perfusion–initial results. *Invest Radiol* 2003; 38: 482–488.
110. Puderbach M, Eichinger M, Gahr J, *et al.* Proton MRI appearance of cystic fibrosis: comparison to CT. *Eur Radiol* 2007; 17: 716–724.
111. Puderbach M, Eichinger M, Haeselbarth J, *et al.* Assessment of morphological MRI for pulmonary changes in cystic fibrosis (CF) patients: comparison to thin-section CT and chest x-ray. *Invest Radiol* 2007; 42: 715–725.
112. Failo R, Wielopolski PA, Tiddens HA, *et al.* Lung morphology assessment using MRI: a robust ultra-short TR/TE 2D steady state free precession sequence used in cystic fibrosis patients. *Magn Reson Med* 2009; 61: 299–306.
113. Eichinger M, Heussel CP, Kauczor HU, *et al.* Computed tomography and magnetic resonance imaging in cystic fibrosis lung disease. *J Magn Reson Imaging* 2010; 32: 1370–1378.
114. Mai VM, Berr SS. MR perfusion imaging of pulmonary parenchyma using pulsed arterial spin labeling techniques: FAIRER and FAIR. *J Magn Reson Imaging* 1999; 9: 483–487.
115. Hatabu H, Gaa J, Kim D, *et al.* Pulmonary perfusion: qualitative assessment with dynamic contrast-enhanced MRI using ultra-short TE and inversion recovery turbo FLASH. *Magn Reson Med* 1996; 36: 503–508.
116. Levin DL, Hatabu H. MR evaluation of pulmonary blood flow. *J Thorac Imaging* 2004; 19: 241–249.
117. Bauman G, Puderbach M, Deimling M, *et al.* Non-contrast-enhanced perfusion and ventilation assessment of the human lung by means of fourier decomposition in proton MRI. *Magn Reson Med* 2009; 62: 656–664.
118. Bauman G, Eichinger M, Deimling M, *et al.* Non-contrast enhanced lung perfusion MRI in comparison to contrast-enahnced MRI perfusion in young cystic fibrosis patients. *Am J Respir Crit Care Med* 2010; 181: Suppl., A1828.
119. Mentore K, Froh DK, de Lange EE, *et al.* Hyperpolarized HHe 3 MRI of the lung in cystic fibrosis: assessment at baseline and after bronchodilator and airway clearance treatment. *Acad Radiol* 2005; 12: 1423–1429.
120. Armstrong DS, Grimwood K, Carzino R, *et al.* Lower respiratory infection and inflammation in infants with newly diagnosed cystic fibrosis. *BMJ* 1995; 310: 1571–1572.
121. de Blic J, Midulla F, Barbato A, *et al.* Bronchoalveolar lavage in children. ERS Task Force on bronchoalveolar lavage in children. European Respiratory Society. *Eur Respir J* 2000; 15: 217–231.
122. Rosenfeld M, Emerson J, Accurso F, *et al.* Diagnostic accuracy of oropharyngeal cultures in infants and young children with cystic fibrosis. *Pediatr Pulmonol* 1999; 28: 321–328.
123. Konstan MW, Berger M. Current understanding of the inflammatory process in cystic fibrosis: onset and etiology. *Pediatr Pulmonol* 1997; 24: 137–142.
124. Elizur A, Cannon CL, Ferkol TW. Airway inflammation in cystic fibrosis. *Chest* 2008; 133: 489–495.
125. Sly PD, Gangell CL, Chen L, *et al.* Risk factors for bronchiectasis in children with cystic fibrosis. *N Engl J Med* 2013; 368: 1963–1970.
126. Radhakrishnan DK, Corey M, Dell SD. Realities of expectorated sputum collection in the pediatric cystic fibrosis clinic. *Arch Pediatr Adolesc Med* 2007; 161: 603–606.
127. Al-Saleh S, Dell SD, Grasemann H, *et al.* Sputum induction in routine clinical care of children with cystic fibrosis. *J Pediatr* 2010; 157: 1006–1011.
128. Blau H, Linnane B, Carzino R, *et al.* Induced sputum compared to bronchoalveolar lavage in young, non-expectorating cystic fibrosis children. *J Cyst Fibros* 2013 [In press DOI: 10.1016/j.jcf.2013.05.013].
129. Mussaffi H, Fireman EM, Mei-Zahav M, *et al.* Induced sputum in the very young: a new key to infection and inflammation. *Chest* 2008; 133: 176–182.
130. Henig NR, Tonelli MR, Pier MV, *et al.* Sputum induction as a research tool for sampling the airways of subjects with cystic fibrosis. *Thorax* 2001; 56: 306–311.
131. De Boeck K, Alifier M, Vandeputte S. Sputum induction in young cystic fibrosis patients. *Eur Respir J* 2000; 16: 91–94.

132. Ordonez CL, Stulbarg M, Grundland H, *et al.* Effect of clarithromycin on airway obstruction and inflammatory markers in induced sputum in cystic fibrosis: a pilot study. *Pediatr Pulmonol* 2001; 32: 29–37.
133. Sagel SD, Kapsner R, Osberg I, *et al.* Airway inflammation in children with cystic fibrosis and healthy children assessed by sputum induction. *Am J Respir Crit Care Med* 2001; 164: 1425–1431.
134. Suri R, Marshall LJ, Wallis C, *et al.* Safety and use of sputum induction in children with cystic fibrosis. *Pediatr Pulmonol* 2003; 35: 309–313.
135. Sagel SD, Sontag MK, Wagener JS, *et al.* Induced sputum inflammatory measures correlate with lung function in children with cystic fibrosis. *J Pediatr* 2002; 141: 811–817.
136. Kim JS, Okamoto K, Rubin BK. Pulmonary function is negatively correlated with sputum inflammatory markers and cough clearability in subjects with cystic fibrosis but not those with chronic bronchitis. *Chest* 2006; 129: 1148–1154.
137. Mayer-Hamblett N, Aitken ML, Accurso FJ, *et al.* Association between pulmonary function and sputum biomarkers in cystic fibrosis. *Am J Respir Crit Care Med* 2007; 175: 822–828.
138. Sagel SD, Wagner BD, Anthony MM, *et al.* Sputum biomarkers of inflammation and lung function decline in children with cystic fibrosis. *Am J Respir Crit Care Med* 2012; 186: 857–865.
139. Liou TG, Adler FR, Keogh RH, *et al.* Sputum biomarkers and the prediction of clinical outcomes in patients with cystic fibrosis. *PLoS One* 2012; 7: e42748.
140. Nick JA, Sanders LA, Ickes B, *et al.* Blood mRNA biomarkers for detection of treatment response in acute pulmonary exacerbations of cystic fibrosis. *Thorax* 2013; 68: 929–937.
141. De Boeck K, Kent L, Davies J, *et al.* CFTR biomarkers: time for promotion to surrogate end-point. *Eur Respir J* 2013; 41: 203–216.
142. Derichs N, Sanz J, Von Kanel T, *et al.* Intestinal current measurement for diagnostic classification of patients with questionable cystic fibrosis: validation and reference data. *Thorax* 2010; 65: 594–599.
143. Veeze HJ, Halley DJ, Bijman J, *et al.* Determinants of mild clinical symptoms in cystic fibrosis patients. Residual chloride secretion measured in rectal biopsies in relation to the genotype. *J Clin Invest* 1994; 93: 461–466.
144. De Jonge HR, Ballmann M, Veeze H, *et al. Ex vivo* CF diagnosis by intestinal current measurements (ICM) in small aperture, circulating Ussing chambers. *J Cyst Fibros* 2004; 3: Suppl. 2, 159–163.
145. Berschneider HM, Knowles MR, Azizkhan RG, *et al.* Altered intestinal chloride transport in cystic fibrosis. *FASEB J* 1988; 2: 2625–2629.
146. Hardcastle J, Hardcastle PT, Taylor CJ, *et al.* Failure of cholinergic stimulation to induce a secretory response from the rectal mucosa in cystic fibrosis. *Gut* 1991; 32: 1035–1039.
147. Hirtz S, Gonska T, Seydewitz HH, *et al.* CFTR Cl^- channel function in native human colon correlates with the genotype and phenotype in cystic fibrosis. *Gastroenterology* 2004; 127: 1085–1095.
148. Mall M, Bleich M, Kuehr J, *et al.* CFTR-mediated inhibition of epithelial Na^+ conductance in human colon is defective in cystic fibrosis. *Am J Physiol* 1999; 277: G709–G716.
149. Mall M, Wissner A, Seydewitz HH, *et al.* Defective cholinergic Cl^- secretion and detection of K^+ secretion in rectal biopsies from cystic fibrosis patients. *Am J Physiol Gastrointest Liver Physiol* 2000; 278: G617–G624.
150. Taylor CJ, Baxter PS, Hardcastle J, *et al.* Failure to induce secretion in jejunal biopsies from children with cystic fibrosis. *Gut* 1988; 29: 957–962.
151. Taylor CJ, Hughes H, Hardcastle PT, *et al.* Genotype and secretory response in cystic fibrosis. *Lancet* 1992; 339: 67–68.
152. Amaral MD, Kunzelmann K. Molecular targeting of CFTR as a therapeutic approach to cystic fibrosis. *Trends Pharmacol Sci* 2007; 28: 334–341.

Chapter 5

Pulmonary arterial hypertension

Gabor Kovacs[*,#] *and Horst Olschewski*[*,#]

SUMMARY: The number of clinical studies on pulmonary arterial hypertension (PAH) and, as a result, the number of available drugs for the disease has grown rapidly over the past two decades. Epoprostenol was the first targeted therapy that was registered for PAH. It was followed by further prostanoids (treprostinil and iloprost), endothelin receptor antagonists (bosentan and ambrisentan) and phosphodiesterase-5 inhibitors (sildenafil and tadalafil). Two additional drugs have also currently been submitted for approval. Further clinical trials have evaluated combination therapy in PAH and PAH drugs in non-PAH pulmonary hypertension. Recently, pilot studies have addressed the effects of PAH drugs in patients with very early pulmonary vascular changes who do not fulfil the criteria of manifest PAH. The aim of this chapter is to provide a guide to the most important previous and current PAH trials in order to give an overview of the current therapeutic approaches in PAH.

*Department of Internal Medicine, Division of Pulmonology, Medical University of Graz, Graz, and #Ludwig Boltzmann Institute for Lung Vascular Research, Graz, Austria.

Correspondence: H. Olschewski, Ludwig Boltzmann Institute for Lung Vascular Research, Stiftingtalstrasse 24, 8010, Graz, Austria.
Email: horst.olschewski@medunigraz.at

Eur Respir Monogr 2013; 62: 70–84.
Copyright ERS 2013.
DOI: 10.1183/1025448x.10044712
Print ISBN: 978-1-84984-044-6
Online ISBN: 978-1-84984-045-3
Print ISSN: 1025-448x
Online ISSN: 2075-6674

In the past two decades, pulmonary hypertension (PH) has become one of the major research areas in respiratory medicine. Numerous clinical trials have been performed and the number of available drugs keeps increasing. Currently, there are almost 2000 clinical trials under the term "pulmonary hypertension" and almost 400 under the term "pulmonary arterial hypertension" available in PubMed (fig. 1). During the past years different end-points have been set in clinical trials with a general shift from simple parameters to composite clinical end-points. The aim of this chapter is to provide a guide to the most influential trials in the treatment of pulmonary arterial hypertension (PAH).

We begin with a discussion on the most frequently used end-points in PAH trials. This will be followed by an overview of the most important clinical trials that have led to the approval of PAH targeted drugs. We will then move on to show the actual data of candidate drugs which may be approved in the near future. This will be followed by a discussion of the trials that have delivered some information on the efficacy of combination therapy *versus* monotherapy in PAH. We will then comment on the most important clinical trials performed in non-PAH PH. Finally, we will discuss clinical trials where PAH drugs were tested in patients with an excessive increase in exercise-induced pulmonary arterial pressure (PAP). Throughout this chapter we will primarily focus on multicentre, randomised, placebo-controlled trials (with a few exceptions), and we aim to show not only the positive trials but also some of the important trials with a negative outcome.

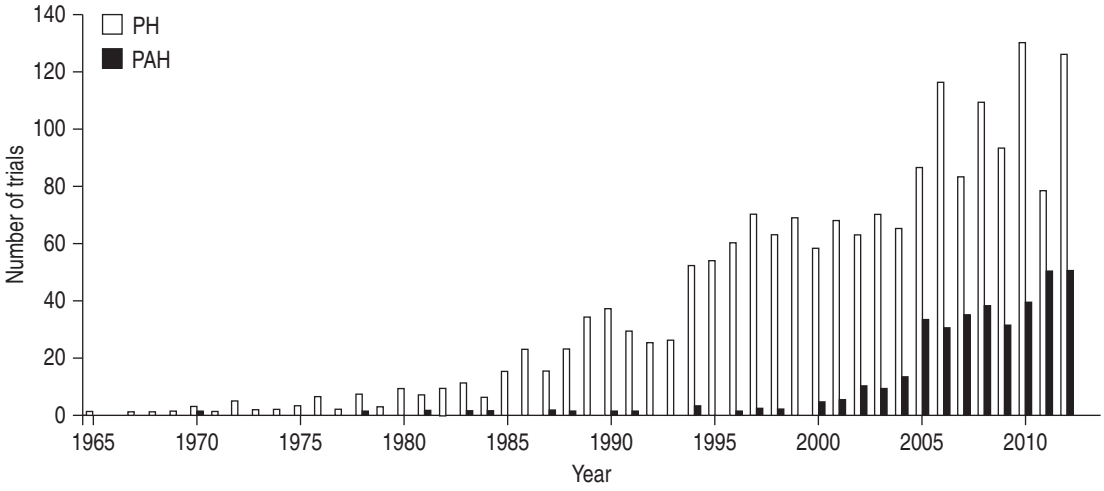


Figure 1. Number of pulmonary hypertension (PH) and pulmonary arterial hypertension (PAH) clinical trials available in PubMed between 1965 and 2012.

End-points of clinical trials in PAH

The establishment of end-points for PH trials has been one of the key issues discussed at the past two World Symposium on Pulmonary Hypertension. It is generally accepted that primary end-points must be clinically relevant, sensitive to treatment effects, and measurable and interpretable [1]. Secondary end-points should provide additional information on the benefit of the tested drug or on the risk-to-benefit ratio.

Exercise capacity

In most previous PH trials, the 6-min walking distance (6MWD) was used as primary end-point. The 6MWD has been shown to be a prognostic indicator [2]. It is simple, reproducible and cheap, and it has been accepted by regulatory agencies. In addition, clinical experience showed that in a medium sized cohort after 3–6 months therapy, a significant therapy effect could be detected by means of the 6MWD. These are important features because PH is an orphan disease, where the inclusion of several thousand patients into clinical trials is merely impossible, and it is a life-threatening disease where long-term placebo-controlled studies may not be ethical. This became obvious after the first epoprostenol trials, where patients in the placebo group experienced rapid deterioration [3]. For this reason, the best of all end-points, all-cause mortality, cannot be employed in patients with severe PAH. As a consequence, nearly all currently available drugs were registered based upon trials using the 6MWD as a primary end-point and a trial length of 3–4 months.

Of course, the 6MWD may be criticised for many reasons. In fact, a recent meta-analysis of 22 PAH trials and 3112 participants detected no significant relationship between the change in 6MWD and clinical outcome parameters. Therefore, an improvement in 6MWD may not necessarily reflect a beneficial clinical outcome [4]. Recent statistical approaches found significant, although weak, associations between the change in 6MWD and the clinical outcomes of PAH patients. Based on a re-analysis of the PHIRST trial, a 33 m improvement in 6MWD was suggested as an estimated minimal important difference [5]. In another meta-analysis, based on the data of 10 PAH trials, a 41.8 m change in 6MWD indicated a statistically significant reduction in clinical events [6]. The question is how these new results can be incorporated in the planning of clinical trials. A further problem with the 6MWD may be the so called ''ceiling effect'' in patients with a 6MWD >450 m [7], and the influence of obvious subjective and objective confounders

independent from PH including musculoskeletal and psychological factors. In spite of these limitations, the 6MWD has been used even in very recent clinical trials as a primary end-point [8, 9].

The peak oxygen uptake during cardiopulmonary exercise testing may be an alternative end-point describing exercise capacity more accurately than the 6MWD. However, this method may be technically challenging to perform and interpret in multicentre trials. Such difficulties were experienced in the STRIDE-1 [10] trial and, since that trial, the method has not been used as primary end-point.

Time to clinical worsening

After the introduction of the first generation of PAH drugs, new first-line drugs can barely be tested against placebo in most parts of the world. However, this limitation is not valid for second-line drugs. Therefore, second- and third-generation drugs were tested as add-on therapy to established first-line drugs. This complicates the comparison of the effects of first-line drugs with later drug generations. In some cases the hurdle appears to be higher [11, 12], but the presence of an effective first-line PAH drug in the placebo arm makes long-term trials possible from an ethical point of view. This also opens the door for choosing clinically relevant end-points such as mortality or time to clinical worsening (TTCW), which have been recommended as primary end-points for future clinical trials [1].

Mortality on its own still appears to be a difficult trial end-point due to the relatively low event rate in patients receiving modern PAH therapy who are stable at inclusion. TTCW may be a surrogate for mortality. It is always a composite end-point; however, different definitions have been employed [1]. It has been recommended that predefined and measurable events such as all-cause mortality, non-elective hospital stay or well-defined disease progression should be included in the definition of TTCW and these events should be controlled by an adjudication committee.

The first successful pivotal trial using TTCW as a primary end-point was the SERAPHIN study which included 742 patients who were followed for 3.5 years [13]. TTCW was defined as death, atrial septostomy, lung transplantation, initiation for *i.v.* or subcutaneous prostanoids, or a combination of the decrease of the 6MWD by 15% confirmed on two different days and worsening functional class or appearance of worsening symptoms of right heart decompensation and therapy escalation. All events were adjudicated by a blinded clinical events committee.

The gap between mortality and TTCW may be closed by a recent analysis of the REVEAL registry, showing that clinical worsening was highly predictive of subsequent proximate mortality. This suggests that clinical worsening is a meaningful prognostic tool in clinical practice and a valid primary end-point in PAH trials [14].

Although definitely more complex than the 6MWD, it is expected that by using TTCW as a composite primary end-point in PAH studies we will receive clinically more relevant information from clinical trials and we may better understand the effects of new drugs before their introduction into clinical practice.

Pivotal clinical trials in PAH leading to the approval of new drugs

Prostanoids

Epoprostenol

Continuously administered *i.v.* epoprostenol was the first targeted therapy for PAH. The first randomised trial with epoprostenol was performed in four referral centres to determine the efficacy of continuous intravenous infusion of the drug in primary PH (now idiopathic and heritable PAH) [15]. The randomised control group was not blinded because a tunnelised central venous catheter and a portable pump for placebo infusion were not utilised. The randomised 8-week treatment period was succeeded by non-randomised treatment for up to 18 months.

Out of the 24 randomised patients, 19 completed the study. In the epoprostenol group there was a strong decrease in pulmonary vascular resistance (PVR; by ~30%) after 8 weeks and six out of 10 patients experienced a decrease in PAP by >10 mmHg, while there was no haemodynamic improvement in the conventional therapy group. In a substantial number of patients, persistent haemodynamic effects could be observed after an 18-month observation period.

In a 12-week prospective, randomised, multicentre open trial, the effects of epoprostenol plus conventional therapy were compared with those of conventional therapy alone in 81 patients with severe primary PH [16]. The trial showed a significant improvement in haemodynamics (mean PAP:improvemed by 8%, PVR improvemed by 21%), exercise capacity (6MWD improved by 47 m in the epoprostenol arm and decreased by 66 m in the conventional therapy arm) and quality of life in the epoprostenol group. Importantly, eight patients died during the study, all of whom had been randomly assigned to conventional therapy. Therefore, it was concluded that as compared with conventional therapy, continuous intravenous infusion of epoprostenol causes symptomatic and haemodynamic improvement, as well as improved survival in patients with severe primary PH.

Based on these data, in November 1995 epoprostenol was approved for the therapy of PAH by the US Food and Drug Administration (FDA). In the following years the drug has also been effectively tested in patients with PAH associated with systemic sclerosis [17]. As a disadvantage, epoprostenol requires a permanent central venous catheter with the associated risk of serious complications such as sepsis, thromboembolism, and right heart decompensation in case of accidental therapy cessation.

Treprostinil

In the future years, further prostanoids were developed for PAH. The major goal of these novel therapies was to achieve similar benefits as epoprostenol with fewer side-effects and a decreased risk of rebound PH upon therapy cessation. Treprostinil is a stable prostacyclin analogue that can be administered by continuous subcutaneous infusion. It was investigated in a 12-week, double-blind, placebo-controlled multicentre trial including 470 patients with PAH (primary or associated with connective tissue disease or congenital systemic-to-pulmonary shunts) [18]. The study revealed a significant improvement in exercise capacity (the difference in median 6MWD changes between the groups was 16 m) in the treprostinil group, while this was unchanged in the placebo arm. Improvement in exercise capacity was greater in the sicker patients and was dose-related, but independent of disease aetiology. In addition, treprostinil significantly improved indices of dyspnoea, signs and symptoms of PH and haemodynamics. The most common side-effect attributed to treprostinil was infusion site pain (85%) leading to premature discontinuation from the study in 8% of patients. It was concluded that treprostinil is an effective treatment for PAH patients with an acceptable safety profile.

Inhalative administration of treprostinil aimed to reduce systemic side-effects and avoid infusion site pain. The TRIUMPH-I study assessed the efficacy and safety of inhaled treprostinil in PAH patients who were receiving either bosentan or sildenafil therapy [19]. This study was performed approximately 10 years after the first targeted drug therapies in PAH. Within this trial, 235 PAH patients with severe symptoms were included and inhaled either treprostinil or placebo four times daily. The primary end-point was peak 6MWD at 12 weeks. The primary end-point (between-treatment median difference in change from baseline in trough 6MWD at week 12) was met (14 m in favour of study drug). In addition, secondary end-points like quality of life measures and N-terminal pro-brain natriuretic peptide (NT-proBNP) improved significantly. However, there were no improvements in time to clinical worsening, Borg dyspnoea score and New York Heart Association (NYHA) functional class. In an open label extension of the study, long-term therapy with inhaled treprostinil demonstrated persistent benefit for PAH patients who remained on therapy for up to 24 months [20].

In addition, oral prostanoid therapy was investigated. In the FREEDOM-M trial, the oral prostacyclin analogue treprostinil diolamine was tested in 249 newly diagnosed PAH patients not

receiving any endothelin receptor antagonist or phosphodiesterase type-5 inhibitor background therapy [12]. The primary end-point was the change from baseline in 6MWD at week 12. Secondary end-points included Borg dyspnoea index, clinical worsening and symptoms of PAH. The primary end-point of the study was reached (treatment effect: 23 m), but there were no significant changes in any secondary end-point. Apart from the typical prostanoid side-effects, oral treprostinil therapy was generally well tolerated. The study recruited most patients in countries with poor access to targeted PAH therapy. This may justify the employment of a placebo treatment arm without background PAH therapy.

In the FREEDOM-C study, a 16-week multicentre, double-blind, placebo-controlled study, the efficacy and safety of oral sustained-release treprostinil was tested in 350 PAH patients (randomised to placebo or oral treprostinil) on a background therapy with an endothelin receptor antagonist and/or phosphodiesterase type-5 inhibitor [11]. The study failed to reach its primary end-point. There was a nonsignificant placebo-corrected improvement in 6MWD (median change was 11 m at week 16; $p=0.07$). The improvement in 6MWD was dependent on the achieved dose of the drug. Significant improvements in secondary end-points such as dyspnoea fatigue index score and combined 6MWD and Borg dyspnoea score were observed. The study discontinuation rate was relatively high in both the treprostinil and the placebo groups (22% and 14%). Based on these studies, treprostinil was approved as subcutaneous infusion by the European Medicines Agency (EMA) and the FDA and can be used as an intravenous infusion in the case of local side-effects. In addition, the FDA approved inhalative treprostinil. Oral treprostinil has not been approved for therapy of PAH.

Beraprost

Beraprost has long been known as an orally active prostacyclin analogue and has been approved in Japan and some Asia-Pacific countries for primary and secondary PH. Beraprost was tested in 130 PAH patients with NYHA functional class II and III [21]. The study reached its primary end-point, the mean change in 6MWD, at week 12 (25 m). Among the secondary end-points, the difference in the mean change of Borg dyspnoea index improved significantly, however, there were no significant changes in cardiopulmonary haemodynamics and NYHA functional class. Drug-related adverse events were common in the titration phase and decreased in the maintenance period.

The study was followed by a second trial assessing the safety and efficacy of beraprost during a 12-month period in 116 PAH patients with World Health Organization (WHO) functional class II or III [22]. The primary end-point was disease progression during the observed period. During this study, the patients treated with beraprost exhibited less evidence of disease progression and an improved 6MWD at 6 months but not at 9 or 12 months. It was concluded that beneficial effects may occur during early phases of treatment but that this effect may not justify the approval of the drug.

Selexipag

A novel approach is selexipag, a non-prostanoid, orally available, selective prostacyclin receptor agonist. It showed favourable haemodynamic changes compared to placebo in a phase II study in patients with symptomatic PAH receiving an endothelin receptor antagonist and/or a phosphodiesterase type-5 inhibitor therapy [23]. The drug is currently being investigated in a phase III multicentre, randomised controlled study (GRIPHON). The results of which are expected in 2014.

Iloprost

Iloprost has long been known as a stable prostacyclin analogue. It was developed as inhalative therapy for PAH. The inhalative approach was aimed at pulmonary and intrapulmonary selective vasodilatation [24, 25]. Inhalative iloprost was tested in 203 severely symptomatic patients with selected forms of severe PAH and chronic thromboembolic pulmonary hypertension (CTEPH) [26]. In this 12-week study, a complex primary end-point was chosen, consisting of the improvement of NYHA class and 6MWD by at least one class and at least 10%, respectively, in the absence of relevant

clinical deterioration or death. The combined clinical end-point was met by 16.8% of the patients receiving iloprost, as compared with 4.9% of the patients on placebo ($p<0.001$). Placebo corrected 6MWD improved by 36.4 m in the whole iloprost group and by 58.8 m in the subgroup of patients with primary PH. Further significant beneficial effects of iloprost treatment included an improvement in haemodynamics, the NYHA class, dyspnoea and quality of life. In a 2-year open-label long-term observation study [27], inhaled iloprost was well tolerated and no substantial dose increase was required. Inhaled iloprost was approved by the EMA and FDA.

In addition to the pivotal trial, studies have been performed investigating iloprost *versus* placebo as add-on therapy to an existing baseline PAH therapy. In the 12-week, randomised, multicentre, double-blind STEP study, the safety and efficacy of adding inhaled iloprost or placebo to bosentan was evaluated in 67 symptomatic patients with PAH [28]. Efficacy end-points included change in 6MWD, modified NYHA functional class, haemodynamic parameters and time to clinical worsening. At week 12, patients receiving iloprost had a mean increase in 6MWD with a placebo-adjusted difference of 26 m ($p=0.051$). In addition, NYHA status, time to clinical worsening, mean PAP and PVR improved significantly in the iloprost group compared to placebo. Combination therapy was well tolerated. Therefore, within the limitations of a relatively small sample size, this study demonstrated that the addition of inhaled iloprost in patients with PAH with reduced exercise capacity on bosentan monotherapy may be safe and efficacious.

Another 12 week, multicentre, open, randomised controlled trial in Germany (COMBI) failed to show positive effects of the addition of inhaled iloprost to bosentan in patients with idiopathic PAH [29]. The study was terminated early due to slow recruitment. None of the secondary end-points including functional class, peak oxygen uptake and time to clinical worsening differed significantly between groups.

Endothelin receptor antagonists

Bosentan

Endothelin receptor antagonists showed beneficial effects in animal models of PH. The dual endothelin receptor antagonist bosentan was developed as the first oral drug for PAH, allowing a simpler administration of PAH therapy. In a phase II study, bosentan was tested in a 12-week double-blind, placebo-controlled preliminary study on 32 patients with PH (primary or associated with scleroderma) [30]. The primary end-point was the change in 6MWD. Secondary end-points included changes in cardiopulmonary haemodynamics, Borg dyspnoea index, WHO functional class, and withdrawal due to clinical worsening. In patients receiving bosentan, as compared to placebo, the distance walked in 6 min improved by 76 m at 12 weeks. Apart from the primary end-point all secondary end-points were significantly improved in the bosentan group compared to placebo.

In a phase III pivotal double-blind, placebo-controlled randomised 16-week trial (BREATHE-1) the effect of bosentan on exercise capacity was investigated [31]. In this trial, 213 patients with PAH (primary or associated with connective-tissue disease) were included. The primary end-point was the change in exercise capacity at 16 weeks. Secondary end-points included the change in Borg dyspnoea index, the change in the WHO functional class, and the time to clinical worsening. At week 16, patients treated with bosentan had an improved 6MWD; the mean difference between the placebo group and the combined bosentan groups (different doses) was 44 m. Bosentan also improved the Borg dyspnoea index and WHO functional class and increased the time to clinical worsening.

In the early clinical trials mainly patients with severe symptoms (NYHA III–IV) were included. The following studies focused on patients with less severe symptoms. In the EARLY study, bosentan was tested in 185 PAH patients with WHO functional status II [32]. The primary end-points of this 6-month double-blind placebo-controlled, multicentre trial were PVR and change in 6MWD. At 6 months, PVR decreased by 17% in the bosentan group and increased by 7% in the placebo group, while 6MWD increased by 11 m in the bosentan group and decreased by

8 m in the placebo group. It was concluded that bosentan treatment is beneficial in patients with WHO functional class II PAH. After this trial the approval of bosentan was extended from WHO functional class III to functional classes II and III.

Due to increased rates of liver enzyme elevations in patients on bosentan, after the approval a large web-based registry was established (TRAcleer eXcellence Post-Marketing Surveillance; Actelion, Allschwil, Switzerland) [33], which allowed the incidence of hepatotoxicity in a large patient population to be assessed. The annualised rate of significant liver enzyme elevations was 10.5%. This necessitates laboratory controls every 4 weeks in all patients on bosentan. There are multiple drug–drug interactions due to effects of bosentan on cytochrome isoenzymes and bile salt transporters.

Sitaxsentan

Sitaxsentan is a selective endothelin-A receptor antagonist for once daily administration due to its long half-life. In the pivotal placebo-controlled 12-week STRIDE-1 study, patients with PAH (idiopathic or related to connective tissue disease or congenital heart disease) were included [10]. Two doses of sitaxsentan (100 and 300 mg) were tested and the primary end-point was the change in peak oxygen uptake ($V'O_2$). Secondary end-points included 6MWD, NYHA class, $V'O_2$ at anaerobic threshold, minute ventilation per carbon dioxide production at the anaerobic threshold, haemodynamics, quality of life and time to clinical worsening. The 300 mg group increased peak $V'O_2$ by just 3% compared with placebo, none of the other end-points derived from cardiopulmonary exercise testing were met. In contrast, both the 100 mg dose and the 300 mg dose improved 6MWD (by 33–35 m), functional class, cardiac index and PVR compared with placebo.

In a second pivotal double-blind, placebo-controlled 18-week study (STRIDE-2), 247 PAH patients (idiopathic or associated with connective tissue disease or congenital heart disease) were included [34]. The primary end-point was change in 6MWD, while secondary end-points included the change in WHO functional class, time to clinical worsening and change in Borg dyspnoea score. At week 18, patients treated with 100 mg of sitaxsentan had an increased 6MWD compared with the placebo group (by 31 m), and an improved WHO functional class. In addition, the incidence of elevated hepatic transaminases during the study was similar in the treatment and placebo group. In the long-term observation study, at 1 year sitaxsentan therapy was considered safe and efficacious for patients with PAH [35]. Sitaxsentan was approved by the EMA for the treatment of PAH. In 2011, sitaxsentan was withdrawn from the market by the manufacturer due to the occurrence of a few cases of idiosyncratic liver failure.

Ambrisentan

Ambrisentan is another selective endothelin-A receptor antagonist for the once-daily treatment of PAH. It was successfully tested in a phase II trial [36] and two pivotal phase III trials [37]. The phase III study programme consisted of the ARIES-1 and ARIES-2 studies. These concurrent, double-blind, placebo-controlled multicentre studies randomised 202 and 192 patients with PAH, respectively, to placebo or different doses (2.5–10 mg once daily) of ambrisentan. The studies were identical in design except for the investigative sites and the doses of ambrisentan studied. ARIES-1 was conducted in 46 centres in the USA, Mexico, South America, Australia and Europe; ARIES-2 was conducted in 41 centres in Europe, Israel, and South America. In the ARIES-1 study 5 and 10 mg doses of ambrisentan were administered and in the ARIES-2 trial doses of 2.5 and 5 mg. The primary end-point for each study was the change in 6MWD from baseline to week 12. In addition, clinical worsening, WHO functional class, Short Form-36 Health Survey score, Borg dyspnoea score and B-type natriuretic peptide plasma concentrations were assessed. The study was positive, 6MWD increased in a dose-dependent manner (31–59 m) in all ambrisentan groups. Improvements in time to clinical worsening, WHO functional class, Short Form-36 score, Borg dyspnoea score and B-type natriuretic peptide were also observed. Importantly, in patients receiving ambrisentan there was no increased incidence of aminotransferase elevations and the improvement in exercise capacity was sustained for at least 48 weeks. The drug was approved by

the EMA and FDA. The long-term extension data for 2 years of ambrisentan treatment demonstrated sustained improvement in exercise capacity and a low risk of clinical worsening and death in patients with PAH [38]. Ambrisentan is generally well tolerated, but the rate of peripheral oedema formation needs special attention. There are no cytochrome and bile transporter interactions. Despite this the authorities requested the establishment of a registry for hepatotoxicity (in analogy to the TRAX registry). The reported rate of significant liver enzyme elevations has been below the expected rate on placebo medication.

Phosphodiesterase-5 inhibitors

Sildenafil

Sildenafil specifically and potently inhibits phosphodiesterase type-5, an enzyme that metabolizes cyclic guanosine-monophosphate, thereby enhancing the cyclic guanosine-monophosphate mediated relaxation and growth inhibition in vascular smooth muscle cells, including those in the lung. Due to the short half-life the medication is administered three times daily. The drug was evaluated in a 12-week double-blind, placebo-controlled study (SUPER-1) in 278 patients with symptomatic PAH [39]. The primary end-point, the 6MWD, increased from baseline in all sildenafil dose groups (by 45–50 m). In addition, sildenafil also reduced the mean PAP and improved WHO functional class. Sildenafil was approved for PAH by the EMA and FDA. In the long-term extension of the study, among the 222 patients completing 1 year of treatment with sildenafil monotherapy, the improvement in 6MWD from baseline at 1 year was maintained. According to the long-term SUPER-2 trial [40], even after 3 years the majority of patients (60%) who entered the SUPER-1 trial improved or maintained their functional status and 46% maintained or improved 6MWD. Sildenafil is generally well tolerated. Heartburn and headache represent the most common side-effects.

Tadalafil

Tadalafil is another phosphodiesterase type-5 inhibitor and presents with a long half-life for once-daily use. This drug was investigated in a 16-week, double-blind, placebo-controlled study (PHIRST) in 405 patients with PAH (either treatment-naïve or on background therapy with bosentan) [41]. Tadalafil increased the 6MWD (primary end-point) in those receiving a dose of 40 mg. Overall, the mean placebo-corrected treatment effect was 33 m (44 m in the treatment naïve group and 23 m in the group on background bosentan therapy). Among the secondary end-points, 40 mg of tadalafil improved the time to clinical worsening and health-related quality of life and was approved for PAH by the EMA and FDA. In a long-term extension study [42], long-term treatment with tadalafil was well tolerated and the improvements in 6MWD demonstrated in the 16-week PHIRST study were sustained for up to 52 additional weeks of treatment.

The PHIRST-1b study separately analysed patients from the PHIRST (tadalafil) study who were either treatment naïve or on background therapy with bosentan, before receiving tadalafil or placebo [43]. At week 16, treatment naïve patients receiving tadalafil had the largest increase in 6MWD (44 m), while patients receiving tadalafil on a background bosentan therapy and placebo on background bosentan therapy had a smaller improvement (23 and 19 m, respectively), while patients on placebo did not experience any improvement of exercise capacity (-3 m). The add-on therapy with tadalafil was well tolerated.

Recent trials on novel drugs for PAH

Riociguat

Riociguat is the first member of a novel class of compounds, the soluble guanylate cyclase (sGC) stimulators. With a dual mode of action, it is synergistic with endogenous nitric oxide (NO) and also directly stimulates sGC, independent of NO availability. Therefore, riociguat may restore the NO-sGC-cGMP pathway. PATENT-1 was a 12-week, double-blind, randomised,

placebo-controlled phase III study investigating the efficacy and safety of riociguat in 443 patients with symptomatic PAH including treatment naïve patients and patients pre-treated with endothelin receptor antagonists or prostanoids [8]. The study reached its primary end-point, which was change in 6MWD (36 m *versus* placebo). This increase was very similar both in treatment naïve and pre-treated patients. In addition, all secondary end-points, including the change from baseline in PVR, NT-proBNP, functional class, clinical worsening, living with PH questionnaire and Borg dyspnoea score, improved significantly. Riociguat was well tolerated and had a favourable safety profile. It has been submitted to the EMA and the FDA for approval for PAH.

Macitentan

Macitentan is another novel drug for PAH. It is an oral dual (endothelin A + endothelin B receptor) endothelin receptor antagonist. Macitentan has a long half-life and is administered once daily. Macitentan was investigated in a large, long-term, event-driven study (SERAPHIN) comparing two different doses of macitentan with placebo in 742 patients with symptomatic PAH [44]. Treatment with macitentan resulted in a reduction in the combined morbidity/mortality end-point with treatment *versus* placebo [13]. This risk was reduced by 45% ($p<0.001$) for patients in the 10 mg dose group and 30% ($p=0.01$) for patients in the 3 mg dose group. Patients in SERAPHIN were treated for up to 3.5 years, providing long-term data showing that macitentan was safe and well tolerated. There are no major cytochrome and bile salt transporter interactions and no indication of hepatotoxicity. Macitentan has been submitted to the EMA and FDA for approval for the therapy of PAH.

Imatinib

In past years, tyrosine kinase inhibitors targeting platelet-derived growth factor were considered as potentially effective new therapy options in PAH. Imatinib is one of these drugs. Imatinib showed promising effects in case reports [45] and a phase II study [46]. The IMPRES study was a randomised, double-blind, placebo-controlled 24-week phase III trial which included 202 symptomatic PAH patients with severely increased PVR who were already on combination therapy for PAH [47]. The primary end-point of the study was change in 6MWD, which was reached (32 m); an effect maintained in the extension study in patients remaining on imatinib. Among the secondary end-points, PVR strongly decreased, but functional class, time to clinical worsening and mortality did not. In addition, serious adverse events and discontinuations were more frequent with imatinib than placebo treatment (44% *versus* 30% and 33% *versus* 18%, respectively). The application for approval has been withdrawn by the sponsor because the EMA scientific committee did not rate the benefit–risk ratio as positive. Subdural haematoma occurred in 10 patients (two in the core study, eight in the open-label extension) on imatinib in combination with anticoagulation, mostly phenprocoumon, out of a population of 241 patients who received imatinib. The use of imatinib for PAH is currently not recommended.

Negative trials

Hydralazine was one of the first drugs used for the treatment of PAH. Despite considerable vasodilative potency it did not improve survival as reliably as high-dose calcium channel blockers [48]. In the small subgroup of patients who acutely respond to inhaled NO with a significant PAP decrease ("responders"), high-dose calcium channel blockers mostly provide an excellent survival benefit [49, 50]. Adenosine is a strong short-acting pulmonary vasodilator. Unfortunately, it is not suitable for long-term therapy [51]. L-arginine supplementation is effective in rat models of PH. A randomised, controlled, double-blind study of L-arginine supplementation in PAH found no improvement compared to placebo (Heart Bar Study Report; unpublished data). Vasoactive intestinal peptide is a strong pulmonary vasodilator with favourable anti-remodelling properties in animal models of PAH. In a randomised double-blind study, inhalation with vasoactive intestinal peptide did not show any

favourable effects as compared to placebo (N. Galie, Dept of Experimental, Diagnostic and Specialty Medicine – DIMES, Bologna University Hospital, Bologna, Italy; personal communication).

Most important clinical trials of PAH medications in non-PAH PH

Left heart disease

In left heart disease PH is quite common and important, as the extent of PH is prognostically relevant in this group of patients. Trials have been performed in order to investigate a potential beneficial effect of PAH drugs; however, until now, none of the multicentre, placebo-controlled double-blind studies were positive.

An early trial in this indication evaluated the effects of epoprostenol on 471 patients with severe left ventricular failure characterised by decreased left ventricular ejection fraction [52]. The primary end-point was survival; secondary end-points were clinical events, congestive heart failure symptoms, 6MWD and quality of life measures. During the trial a significant increase in cardiac index and a decrease in pulmonary capillary wedge pressure were observed. However, the trial was terminated early because of a strong trend toward decreased survival in the patients treated with epoprostenol as compared to placebo.

Among the targeted PAH drugs the phosphodiesterase-5 inhibitor sildenafil probably had the most promising results from pilot trials. Of these, the largest controlled trial included 44 patients from Italy which showed significant effects in favour of sildenafil [53]. In contrast, the RELAX study, a 24-week multicentre, double-blind, placebo-controlled randomised phase III clinical trial was negative. It was performed to test the effects of sildenafil compared with placebo on exercise capacity and clinical status in 216 stable outpatients with heart failure with preserved ejection fraction [54]. Patients did not necessarily present with pulmonary hypertension. The primary end-point, the change in peak $V'O_2$, was not reached. In addition, there was no improvement in the clinical status and the 6MWD. The number of adverse and serious adverse events was slightly (nonsignificant) elevated in the sildenafil group.

Riociguat was tested in 201 patients with PH associated with left ventricular systolic dysfunction in the LEPHT study [55]. This double-blind, phase IIb study characterised the safety and haemodynamic effects of 16 weeks of treatment. The primary end-point of the study (decrease in mean PAP) was not reached. Riociguat, however, resulted in a significant increase in cardiac index and stroke volume index, and in a significant decrease of the pulmonary and systemic vascular resistance. In addition, riociguat improved quality of life and prolonged time to clinical worsening in patients with clinical worsening events (secondary end-points).

Pulmonary disease

In pulmonary diseases, PH is a severe complication, associated with increased mortality. There are some small, mainly monocentre trials with PAH drugs; however, there is a lack of large multicentre, randomised, placebo-controlled clinical trials.

An exception was a randomised, double-blind, placebo-controlled, multicentre study (ARTEMIS-PH) which evaluated the efficacy and safety of ambrisentan in subjects with idiopathic pulmonary fibrosis and PH. The primary end-point of this 16-week trial was the change in 6MWD. The study was terminated early due to lack of benefit. Therefore, as in PH due to left heart diseases, in PH due to pulmonary diseases or hypoxia there is currently no available targeted therapy of PH.

Chronic thromboembolic pulmonary hypertension

CTEPH may result from symptomatic or asymptomatic acute venous thromboembolism. The primary therapy for CTEPH is pulmonary endarterectomy (PEA), but it has long been asked whether those patients with surgically inoperable CTEPH may profit from a targeted PAH therapy.

The first large multicentre study, the BENEFIT trial, was performed to investigate the effect of treatment with bosentan on the haemodynamics and exercise capacity of 157 patients with inoperable disease or persistent/recurrent PH after PEA [56]. There were two independent co-primary end-points, the change in PVR and the change in 6MWD from baseline after 16 weeks of treatment with bosentan or placebo. Although a statistically significant treatment effect of bosentan *versus* placebo was demonstrated (24% decrease in PVR), the mean treatment effect on 6MWD was just 2 m (nonsignificant).

A second large trial for inoperable CTEPH patients has recently been conducted. The CHEST-1 study investigated the efficacy and safety of riociguat in 261 inoperable CTEPH patients. The primary outcome measure was the change in 6MWD from baseline after 16 weeks of treatment. Secondary end-points included change from baseline in PVR, functional class, time to clinical worsening, and safety and tolerability [9]. The CHEST-1 study met its primary end-point by demonstrating a statistically significant improvement in the 6-min walk test (46 m *versus* placebo). Riociguat also showed statistically significant improvements in secondary end-points including PVR, NT-proBNP and WHO functional class. A positive trend was observed in time to clinical worsening and quality of life. Riociguat has been filed for approval for inoperable CTEPH at the EMA and FDA.

Head-to-head comparison of targeted PAH drugs

There are no large, randomised, controlled studies employing a head-to-head comparison of different PAH drugs and it seems unlikely that such a trial will be performed in the near future. There is a small study, however, the 16-week, double-blind, randomised SERAPH study, in which 26 symptomatic PAH patients were randomised to either sildenafil or bosentan and the effects in both groups were compared [57]. Changes in right ventricular mass (using cardiovascular magnetic resonance), 6MWD, cardiac function, brain natriuretic peptide and Borg dyspnoea index were assessed and no significant difference was seen between both groups. Exercise capacity and haemodynamics improved in both groups compared to baseline. Unfortunately, the power of this study was not sufficient to provide definitive answers.

Monotherapy *versus* combination therapy

It has been an important and relevant question as to whether monotherapy followed by sequential combination therapy or an upfront combination therapy is advantageous in PAH patients, however, only a few studies are available.

Currently, the AMBITION study, a randomised, multicentre study, is enrolling patients to answer this question. In this study, the upfront combination of ambrisentan and tadalafil is compared with the individual effects of both drugs. In addition to these much awaited results, there were some studies in which subjects were randomly assigned to placebo or active therapy on a background treatment with an approved PAH drug in a parallel-group design. These studies may give an insight on the effects of combination therapy in PAH.

In the 16-week, double-blind, placebo-controlled prospective BREATHE-2 study [58], the efficacy and safety of combining bosentan and epoprostenol was investigated in 33 PAH patients. All patients received epoprostenol and were randomised to bosentan or placebo. Haemodynamics, exercise capacity and functional class improved in both groups. There was a trend for a greater improvement in all measured haemodynamic parameters but also a larger number of withdrawals in the combination treatment group.

The PACES trial investigated the effect of adding oral sildenafil to intravenous epoprostenol [59]. This 16-week, multicentre, double-blind, placebo-controlled, parallel-group study in 267 PAH patients showed a 28.8 m placebo-adjusted improvement in the 6MWD in patients in the sildenafil group. In addition, regarding secondary end-points, combination with sildenafil resulted

in a greater improvement in mean PAP and cardiac output and an improved time to clinical worsening compared to epoprostenol monotherapy. There was no effect on the Borg dyspnoea score. The addition of sildenafil was generally well tolerated.

Currently, sequential combination therapy is recommended in the international PH guidelines [60], although the evidence was generated from relatively small and partly uncontrolled studies. Upfront combination has been suggested for severely ill patients with right heart decompensation [61].

Exercise-induced PAH

A novel indication for the use of PAH therapy might be the very early form of pulmonary vasculopathy in scleroderma patients. This may be characterised by an excessive PAP increase during exercise and may be of relevance for exercise capacity [62]. Recently, two small trials have been performed in this field which may serve as ''door openers'' in this potential new field of targeted PAH therapy.

The first study examined 10 scleroderma patients with mean PAP >30 mmHg during exercise [63]. All patients underwent right heart catheterisation at baseline after a 12-month observation period and subsequently after 6 months of bosentan therapy. Mean PAP and PVR at rest, at 50 W and during maximal exercise increased significantly during the observation period and this trend was reversed during the treatment period.

A similar study described the changes in haemodynamics and exercise capacity in 12 patients with systemic sclerosis and excessive PAP increase during exercise treated with open-label daily ambrisentan for 24 weeks [64]. There were improvements in mean exercise PVR (primary end-point), mean PAP, mean exercise cardiac output and mean distance covered during the 6MWD.

Before considering the condition ''exercise-induced PAH'' as a therapy indication, large, placebo controlled trials are warranted.

Statement of Interest

G. Kovacs has received speaker's fees from Actelion, Pfizer, GlaxoSmithKline, Boehringer Ingelheim, Bayer, Novartis, Chiesi, AstraZeneca and Takeda. He has also received funding for patents from the Medical University Graz (Graz, Austria) and Siemens. G. Kovacs has also received funding for travel from AOP. H. Olschewski has received a grant from Actelion. He has also received fees for consultancy and participation in review activities from Bayer, Gilead, Lilly and Novartis. H. Olschewski has also received travel grants from Bayer, Gilead, Lilly, Pfizer, GlaxoSmithKline, Actelion and Novartis. He has also received consultancy fees, lecture fees and travel grants from AstraZeneca, GlaxoSmithKline, Pfizer, Almirall and Menarini.

References

1. McLaughlin VV, Badesch DB, Delcroix M, *et al.* End points and clinical trial design in pulmonary arterial hypertension. *J Am Coll Cardiol* 2009; 54: Suppl. 1, S97–S107.
2. Miyamoto S, Nagaya N, Satoh T, *et al.* Clinical correlates and prognostic significance of six-minute walk test in patients with primary pulmonary hypertension. Comparison with cardiopulmonary exercise testing. *Am J Respir Crit Care Med* 2000; 161: 487–492.
3. Barst RJ, Rubin LJ, Long WA, *et al.* A comparison of continuous intravenous epoprostenol (prostacyclin) with conventional therapy for primary pulmonary hypertension. The Primary Pulmonary Hypertension Study Group. *N Engl J Med* 1996; 334: 296–302.
4. Savarese G, Paolillo S, Costanzo P, *et al.* Do changes of 6-minute walk distance predict clinical events in patients with pulmonary arterial hypertension? A meta-analysis of 22 randomized trials. *J Am Coll Cardiol* 2012; 60: 1192–1201.
5. Mathai SC, Puhan MA, Lam D, *et al.* The minimal important difference in the 6-minute walk test for patients with pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2012; 186: 428–433.
6. Gabler NB, French B, Strom BL, *et al.* Validation of 6-minute walk distance as a surrogate end point in pulmonary arterial hypertension trials. *Circulation* 2012; 126: 349–356.

7. Frost AE, Langleben D, Oudiz R, *et al.* The 6-min walk test (6MW) as an efficacy endpoint in pulmonary arterial hypertension clinical trials: demonstration of a ceiling effect. *Vascul Pharmacol* 2005; 43: 36–39.
8. Ghofrani HA, Galie N, Grimminger F, *et al.* Riociguat for the treatment of pulmonary arterial hypertension. *N Engl J Med* 2013; 369: 330–340.
9. Ghofrani HA, D'Armini AM, Grimminger F, *et al.* Riociguat for the treatment of chronic thromboembolic pulmonary hypertension. *N Engl J Med* 2013; 369: 319–329.
10. Barst RJ, Langleben D, Frost A, *et al.* Sitaxsentan therapy for pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2004; 169: 441–447.
11. Tapson VF, Torres F, Kermeen F, *et al.* Oral treprostinil for the treatment of pulmonary arterial hypertension in patients on background endothelin receptor antagonist and/or phosphodiesterase type 5 inhibitor therapy (The FREEDOM-C Study): a randomized controlled trial. *Chest* 2012; 142: 1383–1390.
12. Jing ZC, Parikh K, Pulido T, *et al.* Efficacy and safety of oral treprostinil monotherapy for the treatment of pulmonary arterial hypertension: a randomized, controlled trial. *Circulation* 2013; 127: 624–633.
13. Channick R, Delcroix M, Galie N, *et al.* Macitentan reduces PAH-related hospitalizations: results from the randomized controlled SERAPHIN trial. *Am J Respir Crit Care Med* 2013; 187: A3527.
14. Frost AE, Badesch DB, Miller DP, *et al.* Evaluation of the predictive value of a clinical worsening definition using 2-year outcomes in patients with pulmonary arterial hypertension: a REVEAL registry analysis. *Chest* 2013 [In press DOI: 10.1378/chest.12-3023].
15. Rubin LJ, Mendoza J, Hood M, *et al.* Treatment of primary pulmonary hypertension with continuous intravenous prostacyclin (epoprostenol). Results of a randomized trial. *Ann Intern Med* 1990; 112: 485–491.
16. Barst RJ, Rubin LJ, Long WA, *et al.* A comparison of continuous intravenous epoprostenol (prostacyclin) with conventional therapy for primary pulmonary hypertension. The Primary Pulmonary Hypertension Study Group. *N Engl J Med* 1996; 334: 296–302.
17. Badesch DB, Tapson VF, McGoon MD, *et al.* Continuous intravenous epoprostenol for pulmonary hypertension due to the scleroderma spectrum of disease. A randomized, controlled trial. *Ann Intern Med* 2000; 132: 425–434.
18. Simonneau G, Barst RJ, Galie N, *et al.* Continuous subcutaneous infusion of treprostinil, a prostacyclin analogue, in patients with pulmonary arterial hypertension: a double-blind, randomized, placebo-controlled trial. *Am J Respir Crit Care Med* 2002; 165: 800–804.
19. McLaughlin VV, Benza RL, Rubin LJ, *et al.* Addition of inhaled treprostinil to oral therapy for pulmonary arterial hypertension: a randomized controlled clinical trial. *J Am Coll Cardiol* 2010; 55: 1915–1922.
20. Benza RL, Seeger W, McLaughlin VV, *et al.* Long-term effects of inhaled treprostinil in patients with pulmonary arterial hypertension: the Treprostinil Sodium Inhalation Used in the Management of Pulmonary Arterial Hypertension (TRIUMPH) study open-label extension. *J Heart Lung Transplant* 2011; 30: 1327–1333.
21. Galie N, Humbert M, Vachiery JL, *et al.* Effects of beraprost sodium, an oral prostacyclin analogue, in patients with pulmonary arterial hypertension: a randomized, double-blind, placebo-controlled trial. *J Am Coll Cardiol* 2002; 39: 1496–1502.
22. Barst RJ, McGoon M, McLaughlin V, *et al.* Beraprost therapy for pulmonary arterial hypertension. *J Am Coll Cardiol* 2003; 41: 2119–2125.
23. Simonneau G, Torbicki A, Hoeper MM, *et al.* Selexipag, an oral, selective prostacyclin receptor agonist for the treatment of pulmonary arterial hypertension. *Eur Respir J* 2012; 40: 874–880.
24. Olschewski H, Walmrath D, Schermuly R, *et al.* Aerosolized prostacyclin and iloprost in severe pulmonary hypertension. *Ann Intern Med* 1996; 124: 820–824.
25. Olschewski H, Ghofrani HA, Walmrath D, *et al.* Inhaled prostacyclin and iloprost in severe pulmonary hypertension secondary to lung fibrosis. *Am J Respir Crit Care Med* 1999; 160: 600–607.
26. Olschewski H, Simonneau G, Galie N, *et al.* Inhaled iloprost for severe pulmonary hypertension. *N Engl J Med* 2002; 347: 322–329.
27. Olschewski H, Hoeper MM, Behr J, *et al.* Long-term therapy with inhaled iloprost in patients with pulmonary hypertension. *Respir Med* 2010; 104: 731–740.
28. McLaughlin VV, Oudiz RJ, Frost A, *et al.* Randomized study of adding inhaled iloprost to existing bosentan in pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2006; 174: 1257–1263.
29. Hoeper MM, Leuchte H, Halank M, *et al.* Combining inhaled iloprost with bosentan in patients with idiopathic pulmonary arterial hypertension. *Eur Respir J* 2006; 28: 691–694.
30. Channick RN, Simonneau G, Sitbon O, *et al.* Effects of the dual endothelin-receptor antagonist bosentan in patients with pulmonary hypertension: a randomised placebo-controlled study. *Lancet* 2001; 358: 1119–1123.
31. Rubin LJ, Badesch DB, Barst RJ, *et al.* Bosentan therapy for pulmonary arterial hypertension. *N Engl J Med* 2002; 346: 896–903.
32. Galie N, Rubin L, Hoeper M, *et al.* Treatment of patients with mildly symptomatic pulmonary arterial hypertension with bosentan (EARLY study): a double-blind, randomised controlled trial. *Lancet* 2008; 371: 2093–2100.
33. Humbert M, Segal ES, Kiely DG, *et al.* Results of European post-marketing surveillance of bosentan in pulmonary hypertension. *Eur Respir J* 2007; 30: 338–344.

34. Barst RJ, Langleben D, Badesch D, *et al.* Treatment of pulmonary arterial hypertension with the selective endothelin-A receptor antagonist sitaxsentan. *J Am Coll Cardiol* 2006; 47: 2049–2056.
35. Benza RL, Barst RJ, Galie N, *et al.* Sitaxsentan for the treatment of pulmonary arterial hypertension: a 1-year, prospective, open-label observation of outcome and survival. *Chest* 2008; 134: 775–782.
36. Galie N, Badesch D, Oudiz R, *et al.* Ambrisentan therapy for pulmonary arterial hypertension. *J Am Coll Cardiol* 2005; 46: 529–535.
37. Galie N, Olschewski H, Oudiz RJ, *et al.* Ambrisentan for the treatment of pulmonary arterial hypertension: results of the ambrisentan in pulmonary arterial hypertension, randomized, double-blind, placebo-controlled, multicenter, efficacy (ARIES) study 1 and 2. *Circulation* 2008; 117: 3010–3019.
38. Oudiz RJ, Galie N, Olschewski H, *et al.* Long-term ambrisentan therapy for the treatment of pulmonary arterial hypertension. *J Am Coll Cardiol* 2009; 54: 1971–1981.
39. Galie N, Ghofrani HA, Torbicki A, *et al.* Sildenafil citrate therapy for pulmonary arterial hypertension. *N Engl J Med* 2005; 353: 2148–2157.
40. Rubin LJ, Badesch DB, Fleming TR, *et al.* Long-term treatment with sildenafil citrate in pulmonary arterial hypertension: the SUPER-2 study. *Chest* 2011; 140: 1274–1283.
41. Galie N, Brundage BH, Ghofrani HA, *et al.* Tadalafil therapy for pulmonary arterial hypertension. *Circulation* 2009; 119: 2894–2903.
42. Oudiz RJ, Brundage BH, Galie N, *et al.* Tadalafil for the treatment of pulmonary arterial hypertension: a double-blind 52-week uncontrolled extension study. *J Am Coll Cardiol* 2012; 60: 768–774.
43. Barst RJ, Oudiz RJ, Beardsworth A, *et al.* Tadalafil monotherapy and as add-on to background bosentan in patients with pulmonary arterial hypertension. *J Heart Lung Transplant* 2011; 30: 632–643.
44. Pulido T, Adzerikho I, Channick RN, *et al.* Macitentan and morbidity and mortality in pulmonary arterial hypertension. *N Engl J Med* 2013; 369: 809–818.
45. Ghofrani HA, Seeger W, Grimminger F. Imatinib for the treatment of pulmonary arterial hypertension. *N Engl J Med* 2005; 353: 1412–1413.
46. Ghofrani HA, Morrell NW, Hoeper MM, *et al.* Imatinib in pulmonary arterial hypertension patients with inadequate response to established therapy. *Am J Respir Crit Care Med* 2010; 182: 1171–1177.
47. Hoeper MM, Barst RJ, Bourge RC, *et al.* Imatinib mesylate as add-on therapy for pulmonary arterial hypertension: results of the randomized IMPRES study. *Circulation* 2013; 127: 1128–1138.
48. Sandoval J, Bauerle O, Palomar A, *et al.* Survival in primary pulmonary hypertension. Validation of a prognostic equation. *Circulation* 1994; 89: 1733–1744.
49. Rich S, Kaufmann E, Levy PS. The effect of high doses of calcium-channel blockers on survival in primary pulmonary hypertension. *N Engl J Med* 1992; 327: 76–81.
50. Sitbon O, Humbert M, Jais X, *et al.* Long-term response to calcium channel blockers in idiopathic pulmonary arterial hypertension. *Circulation* 2005; 111: 3105–3111.
51. Haywood GA, Sneddon JF, Bashir Y, *et al.* Adenosine infusion for the reversal of pulmonary vasoconstriction in biventricular failure. A good test but a poor therapy. *Circulation* 1992; 86: 896–902.
52. Califf RM, Adams KF, McKenna WJ, *et al.* A randomized controlled trial of epoprostenol therapy for severe congestive heart failure: the Flolan International Randomized Survival Trial (FIRST). *Am Heart J* 1997; 134: 44–54.
53. Guazzi M, Vicenzi M, Arena R, *et al.* Pulmonary hypertension in heart failure with preserved ejection fraction: a target of phosphodiesterase-5 inhibition in a 1-year study. *Circulation* 2011; 124: 164–174.
54. Redfield MM, Chen HH, Borlaug BA, *et al.* Effect of phosphodiesterase-5 inhibition on exercise capacity and clinical status in heart failure with preserved ejection fraction: a randomized clinical trial. *JAMA* 2013; 309: 1268–1277.
55. Bonderman D, Ghio S, Felix SB, *et al.* Riociguat for patients with pulmonary hypertension caused by systolic left ventricular dysfunction: a phase IIb double-blind, randomized, placebo-controlled, dose-ranging hemodynamic study. *Circulation* 2013; 128: 502–511.
56. Jais X, D'Armini AM, Jansa P, *et al.* Bosentan for treatment of inoperable chronic thromboembolic pulmonary hypertension: BENEFiT (Bosentan Effects in iNopErable Forms of chronIc Thromboembolic pulmonary hypertension), a randomized, placebo-controlled trial. *J Am Coll Cardiol* 2008; 52: 2127–2134.
57. Wilkins MR, Paul GA, Strange JW, *et al.* Sildenafil *versus* Endothelin Receptor Antagonist for Pulmonary Hypertension (SERAPH) study. *Am J Respir Crit Care Med* 2005; 171: 1292–1297.
58. Humbert M, Barst RJ, Robbins IM, *et al.* Combination of bosentan with epoprostenol in pulmonary arterial hypertension: BREATHE-2. *Eur Respir J* 2004; 24: 353–359.
59. Simonneau G, Rubin LJ, Galie N, *et al.* Addition of sildenafil to long-term intravenous epoprostenol therapy in patients with pulmonary arterial hypertension: a randomized trial. *Ann Intern Med* 2008; 149: 521–530.
60. Galie N, Hoeper MM, Humbert M, *et al.* Guidelines for the diagnosis and treatment of pulmonary hypertension. *Eur Respir J* 2009; 34: 1219–1263.
61. Ghofrani HA, Distler O, Gerhardt F, *et al.* Treatment of pulmonary arterial hypertension (PAH): updated recommendations of the Cologne Consensus Conference 2011. *Int J Cardiol* 2011; 154: Suppl. 1, S20–S33.

62. Kovacs G, Maier R, Aberer E, *et al.* Borderline pulmonary arterial pressure is associated with decreased exercise capacity in scleroderma. *Am J Respir Crit Care Med* 2009; 180: 881–886.
63. Kovacs G, Maier R, Aberer E, *et al.* Pulmonary arterial hypertension therapy may be safe and effective in patients with systemic sclerosis and borderline pulmonary artery pressure. *Arthritis Rheum* 2012; 64: 1257–1262.
64. Saggar R, Khanna D, Shapiro S, *et al.* Brief report: effect of ambrisentan treatment on exercise-induced pulmonary hypertension in systemic sclerosis: a prospective single-center, open-label pilot study. *Arthritis Rheum* 2012; 64: 4072–4077.

Chapter 6

The intensive care unit

Donald M. Arnold[*,#] *and Deborah J. Cook*[*,¶]

SUMMARY: In this chapter, we discuss some pros and cons of different selected outcome measures for randomised clinical trials in the intensive care unit. The most important principle is that the outcome measure selected should reflect the objective of the trial. Choosing the best outcome for a randomised controlled trial is one of the most important elements of study design for those running the trial. Interpreting that outcome in the light of other study outcomes is one of the most important skills of readers. Pilot trials are useful tools to ensure mechanistic plausibility and feasibility of larger trials, along with other objectives. While mortality may be the most robust patient-important outcome, other outcomes may be suitable, such as ventilator-free days and quality-adjusted life-years, and could reduce the sample size requirements. Adjudication is a useful tool to ensure validity of outcomes prone to bias in judgement. Ultimately, the choices of primary, secondary and tertiary outcomes need to be informed by the research question and selected in the context of research resources.

*Dept of Medicine, McMaster University, Hamilton,
[#]Canadian Blood Services, Hamilton, and
[¶]Dept of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, ON, Canada.

Correspondence: D.J. Cook, Room D176 St Joseph's Healthcare, 50 Charlton Avenue East, Hamilton, ON, Canada.
Email: debcook@mcmaster.ca

Eur Respir Monogr 2013; 62: 85–89.
Copyright ERS 2013.
DOI: 10.1183/1025448x.10044912
Print ISBN: 978-1-84984-044-6
Online ISBN: 978-1-84984-045-3
Print ISSN: 1025-448x
Online ISSN: 2075-6674

Randomised controlled trials (RCTs) in the intensive care unit (ICU) are crucial to minimise risk of bias in testing many established and new interventions. RCTs require a large investment of time, effort and resources, and require careful planning. One of the most important aspects of RCT design is the choice of the outcome, which should be patient-important and clinically meaningful, and of course relevant to the research question. Preparatory work is usually advisable in the form of pilot trials to determine appropriate outcomes based on mechanisms or feasibility. Ultimately, valid, reproducible and unbiased patient-important outcomes are required for large RCTs.

Pilot trials

Pilot trials are important to ensure that future large RCTs are rigorous, feasible and economically justifiable [1]. The outcomes of pilot RCTs vary depending on the objective of the pilot trial. Phase I and II studies are common terms used to describe pilot studies of drugs or devices. Phase I studies are the first investigations in humans (typically healthy volunteers or uncomplicated patients), designed to evaluate the effect of a new intervention on surrogate efficacy outcomes (*e.g.* biomarkers) or specific harms. Phase II studies are designed to identify therapies or dosing regimens that warrant further investigation based on acceptable toxicity and promising efficacy; if randomised, they can also represent pilot trials. Phase II RCTs are useful when data from historical controls are lacking. Phase III studies are RCTs designed to compare the efficacy of a new intervention with an alternate control (usually the current standard of care), once an intervention

is demonstrated to be reasonably safe and potentially effective. Clearly these different designs lend themselves to different outcome selection.

Two key end-points for pilot RCTs are an assessment of feasibility and a means of understanding mechanisms based on surrogate outcome measures. In this chapter, we outline issues related to feasibility and mechanistic outcomes in pilot RCTs.

Feasibility outcomes in pilot RCTs

The feasibility of conducting a large RCT can be preliminarily tested in a pilot RCT that simulates all aspects of the trial, from the allocation process to recruitment, retention, outcome measurement, data collection and follow-up. A pilot trial can provide important insight into the challenges and determinants of protocol adherence for subsequent large RCTs. For example, extensive preparatory work from Australia and Canada helped to assess the feasibility of implementing an intensive insulin algorithm to achieve two different target glucose ranges, enabling investigators to refine algorithms for glycaemic targets for the large international NICE-SUGAR Study (Normoglycemia in Intensive Care Evaluation – Survival Using Glucose Algorithm Regulation) [2]. Indeed, two pilot RCTs augmented several other nonrandomised pilot studies, preceding implementation of this international trial of intensive insulin therapy in the ICU [2]. A pilot trial conducted at multiple sites helps to ensure that the protocol can be implemented outside experienced, well-resourced centres.

Mechanistic outcomes in pilot trials

Pilot trials sometimes use surrogate outcomes to understand mechanisms before proceeding to a larger RCT. For example, in the Beta Agonist Lung Injury (BALTI) trial, extravascular lung water was measured, and the researchers running the trial concluded that there was sufficient proof of principle to justify a larger trial powered on clinically important outcomes [3]. Similarly, a single-centre pilot trial was designed to assess the safety of small tidal volume ventilation for patients with acute lung injury [4], using ventilatory and haemodynamic parameters as surrogates. A subsequent large multicentre RCT demonstrated that small tidal volume ventilation lowered mortality and increased ventilator-free days (VFD) [5]. Analyses based on surrogate outcomes should be interpreted with caution, keeping in mind that pilot trials are underpowered to detect clinically important treatment effects. Also, pilot trials are prone to the play of chance, due to their typically modest sample size, potentially leading readers to factitiously conclude significant differences when none really exist.

Mortality

The classic outcome for RCTs in the ICU is short-term mortality at 28 days or at ICU discharge. Mortality is patient-important and clearly unbiased. Given the complexity of critical illness and the lack of a universal, agreed-upon approach to accurately assess cause of death in the ICU, cause-specific mortality is rarely a primary outcome of an ICU trial, although it is sometimes a tertiary or secondary outcome. The disadvantages of a short-term horizon for mortality are that the cut-off dates are either arbitrary (*e.g.* 30 days) or purely conventional (*e.g.* 28 days for US Food and Drug Administration regulations). Deaths occurring soon after the cut-off are ignored using such a metric. However, ICU mortality remains a nearly universal outcome, even if not the primary outcome, for ICU RCTs.

Longer-term evaluation of mortality is increasingly recognised as worthwhile. One important recent example of how short-term mortality can be misleading is found in the NICE-SUGAR Study [2]. Of the 6104 randomised patients, 3054 were assigned to intensive control and 3050 to conventional control. There were 829 (27.5%) deaths in the intensive group and 751 (24.9%) deaths in the conventional group at 90 days (OR 1.14, 95% CI 1.02–1.28; p=0.02). However,

this difference in mortality was not evident at 28 days (670 (22.3%) deaths *versus* 627 (20.8%) deaths; OR 1.09, 95% CI 0.96–1.23; p=0.17). Had the primary outcome only been measured at 28 days and had patients not been followed for a full 3 months, the potential harm associated with targeting euglycaemia in the ICU would not have been recognised, and harmful management might have continued for decades.

Ventilator-free days

Trialists face multiple challenges to make an impact on mortality in conditions such as sepsis and acute respiratory distress syndrome (ARDS). Treatment effects in medicine tend to be small, and demonstrating a mortality advantage is difficult. Recently, a new outcome was proposed for RCTs in the ICU: VFD, defined as days alive and free from mechanical ventilation.

A study was conducted to examine the utility, behaviour and statistical properties of this outcome [6] by evaluating two recent RCTs in acute lung injury [5, 7]. Through modelling, investigators found that use of VFD as a trial end-point allows smaller sample sizes if it is assumed that the treatment being tested simultaneously reduces the duration of ventilation and improves mortality. It is unlikely that a treatment that led to higher mortality could lead to a statistically significant improvement in VFD. This would be especially true if the treatment were also required to produce a nominal improvement in mortality [6]. The end-point of VFD is not a very clinically intuitive outcome, and may be influenced heavily by approaches to mechanical ventilation and weaning. To date, few trials in critical care have used this as primary outcome, but several trials have used it as a secondary outcome.

Quality-adjusted life-years

Quality-adjusted life-years (QALYs) are composite outcomes that combine mortality and measures of health-related quality of life. Utilities are determined by patient ratings of health states on a scale of 0 to 1, and 1 represents ''perfect health''. QALYs allow comparisons of outcomes across various diseases and within patients over time, or across arms of an RCT. Because QALYs are a continuous outcome measure, statistical power may be increased when planning a trial if QALYs are used. In a modelling exercise to describe the distribution of QALYs in a cohort of ICU survivors [8], sample size requirements for a hypothetical trial using QALYs as the primary outcome were much lower than if mortality had been used, a relationship which was maintained under many assumptions [9]. However, this approach was susceptible to loss to follow-up, which is common when following ICU survivors after ICU discharge. Another obstacle to using QALYs in RCTs is the associated trial costs for extended contact with ICU survivors.

HERRIDGE *et al.* [10] have pioneered the follow-up of ICU survivors in their studies. For example, extending their original work, in-person interviews and examinations were conducted of 109 survivors of ARDS to obtain follow-up data at 3, 6 and 12 months and at 2, 3, 4 and 5 years after ICU discharge [10]. The median 6-min walking distance at 5 years was 76% of predicted distance. Exercise limitation, physical and psychological sequelae, and decreased physical quality of life were documented. While younger patients had faster recovery than older patients, neither group returned to normal predicted levels of physical function at 5 years and several other physical and psychological problems developed or persisted. Patients with more coexisting illnesses incurred greater 5-year costs to the healthcare system. In addition, family members had psychological dysfunction after their loved one's ICU stay. This work has spawned interest in ICU-acquired weakness, and evaluating the effects of customised, family-centred, rehabilitation programmes on long-term outcomes after a critical illness. Many other measures of cognitive, emotional, psychological, sexual and functional outcomes are now included in RCTs in the ICU, as suited to the research question.

Morbidity outcomes and their adjudication

The measurement of major morbidities in clinical trials is prone to random and systematic error, even if such outcomes are the most suitable for the trial objective. In a recent thromboprophylaxis RCT, known as PROTECT (PROphylaxis for ThromboEmbolism in Critical Care Trial) [11], dalteparin was compared with unfractionated heparin in 3764 medical-surgical ICU patients. The primary outcome was proximal leg deep vein thrombosis (DVT) and secondary end-points were bleeding, pulmonary embolism, DVT at other sites (*e.g.* upper limb, distal leg DVT), and any venous thromboembolism. Since there is a component of subjectivity to each of these outcomes, they were adjudicated. Adjudication can improve the rigour and face validity of such outcomes.

The example of bleeding will be used here as an illustration [12]. In PROTECT, bleeding data were collected prospectively each day using a validated tool. First, an initial calibration exercise was conducted among all bleed adjudicators. For the first 45 bleeding events that occurred in 40 patients, crude agreement between pairs of adjudicators ranged from 86.7% to 93.3% and all four adjudicators agreed on the assessment of bleeding in 82.2% events. Overall, for all four adjudicators, $\kappa=0.81$. As a result of discussions during the calibration phase, the methods and logistics for evaluating bleeding were refined. Thus, after independent quadruplicate review of 45 bleeding events in this calibration exercise, we documented satisfactory agreement for bleeding outcomes, refined the adjudication process and calibrated adjudicators for the remainder of the trial.

Subsequently, the severity of each bleed was adjudicated independently by the principal investigator and one out of three randomly selected adjudicators, blinded to treatment allocation, centre and each other's assessment. Disagreements between pairs of adjudicators were resolved by discussion and consensus. Out of 491 patients with bleeding events in PROTECT, 446 were adjudicated in duplicate after the calibration exercise. Ultimately, 182 (40.8%) patients were classified as having major bleeding, 250 (56.1%) were classified as having minor bleeding only, and 14 (2.9%) were classified as having no bleeding event. Overall agreement among adjudicators was excellent (crude agreement 86.3%; $\kappa=0.76$). Pair-wise agreement between adjudicators was also good ($\kappa=0.87$, 0.75 and 0.69 for each adjudicator pair). Some bleeds were downgraded and others upgraded by adjudicators compared with initial assessments of research coordinators; however, trial results were similar using adjudicated and non-adjudicated events. Central adjudication minimised random and systematic error, improved face validity in the assessment of this major safety outcome, and helped to ensure uniform application of the definitions, standardising assessments across adjudication pairs, across centres and over time.

In an ICU trial that compared the effect of two drugs on the prevention of stress ulcer bleeding, nosocomial pneumonia was a key safety outcome. Pneumonia rates were markedly different depending on whether they were based on submitted data from participating centres, "standard" definitions, professionally recommended definitions or adjudicated assessments [13]. Furthermore, misclassification of even a small number of events can bias trial results especially when outcomes are not common. The added value of outcome adjudication exercises may vary depending upon the objectivity of criteria defining each outcome, the number of outcomes, competing priorities of trial management and resources, the results themselves, and the perceived importance of the outcome to the ultimate interpretation of the trial results.

Other outcomes

The foregoing is by no means an exhaustive list of potential outcomes for ICU trials. Length of ICU stay is another outcome rarely used today as a primary outcome. This is in part dependent on systems issues such as ward bed availability for the timing of transferring patients. Length of hospital stay and other economic outcomes are also common, but not discussed further here.

Other study designs

Clearly, RCTs are not the best study design for all research questions. In critical care, they have been most commonly used to compare preventive or therapeutic interventions and diagnostic tests. Other observational designs, such as case–control studies, may be useful to study the impact of uncommon conditions or exposures that cannot be randomised. Cohort studies may be most suitable to study population-based incidence rates if RCTs are not representative of unselected patients. Aside from quantitative studies, the qualitative research world offers a range of other study designs. Regardless, the outcomes should depend on the research question.

Support Statement

D.J. Cook is a Research Chair of the Canadian Institutes of Health Research.

Statement of Interest

D.M. Arnold holds a New Investigator Award from the Canadian Institutes for Health Research in partnership with Hoffmann-LaRoche.

References

1. Arnold DA, Burns KE, Adhikari N, *et al.* The design and interpretation of pilot trials in critical care medicine. *Crit Care Med* 2009; 36: S69–S74.
2. Finfer S, Chittock V, Yu-Shuo Su S, *et al.* Intensive *versus* conventional glucose control in critically ill patients. *N Engl J Med* 2009; 360: 1283–1297.
3. Perkins GD, McAuley DF, Thickett DR, *et al.* The beta-agonist lung injury trial (BALTI): a randomized placebo-controlled clinical trial. *Am J Respir Crit Care Med* 2006; 173: 281–287.
4. Brower RG, Shanholtz CB, Fessler HE, *et al.* Prospective, randomized, controlled clinical trial comparing traditional *versus* reduced tidal volume ventilation in acute respiratory distress syndrome patients. *Crit Care Med* 1999; 27: 1492–1498.
5. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. The Acute Respiratory Distress Syndrome Network. *N Engl J Med* 2000; 342: 1301–1308.
6. Schoenfeld DA, Bernard GR, ARDS Network. Statistical evaluation of ventilator-free days as an efficacy measure in clinical trials of treatments for acute respiratory distress syndrome. *Crit Care Med* 2002; 30: 1772–1777.
7. Ketoconazole for early treatment of acute lung injury and acute respiratory distress syndrome: a randomized controlled trial. The ARDS Network. *JAMA* 2000; 283: 1995–2002.
8. Herridge MS, Cheung AM, Tansey CM, *et al.* One-year outcomes in survivors of the acute respiratory distress syndrome. *N Engl J Med* 2003; 348: 683–693.
9. Ferguson ND, Scales DC, Pinto R, *et al.* Integrating mortality and morbidity outcomes: using quality-adjusted life years in critical care trials. *Am J Respir Crit Care Med* 2013; 187: 256–261.
10. Herridge MS, Tansey CM, Matte A, *et al.* Functional disability 5 years after acute respiratory distress syndrome. *N Engl J Med* 2011; 364: 1293–1304.
11. PROTECT Investigators for the Canadian Critical Care Trials Group and the Australian and New Zealand Intensive Care Society Clinical Trials Group. Dalteparin *versus* unfractionated heparin in critically ill patients. *N Engl J Med* 2011; 364: 1305–1314.
12. Arnold DM, Lauzier F, Rabbat C, *et al.* Adjudication of bleeding outcomes in an international thromboprophylaxis trial in critical illness. *Thromb Res* 2013; 131: 204–209.
13. Cook DJ, Walter SD, Freitag A, *et al.* Adjudicating ventilator-associated pneumonia in a randomized trial of critically ill patients. *J Crit Care* 1998; 13: 159–163.

Chapter 7

Biomarkers

Jonas Schupp and Antje Prasse

SUMMARY: In the last decade, major advances have been made in the field of biomarker research in various lung diseases, including asthma, chronic obstructive pulmonary disease, lower respiratory tract infections, lung cancer and interstitial lung diseases. Multiple biomarkers have been implemented in the clinical practice of respiratory physicians. In parallel with the evolving field of personalised medicine, the number of clinical trials that utilise biomarkers is increasing every year. The concepts of biomarker-stratified patient selection and targeted therapy have been established and their efficiency successfully proven. In lung cancer research, which takes a pioneering role, biomarker-driven research has become an integral part of nearly every clinical trial. Moreover, biomarkers are used to predict and monitor treatment responses. This chapter summarises current achievements of clinical trials in biomarker research related to lung diseases.

Dept of Pneumology, University Medical Clinic Freiburg, Freiburg, Germany.

Correspondence: A. Prasse, Dept of Pneumology, University Medical Clinic Freiburg, Kilianstrasse 5, 79106 Freiburg, Germany.
Email: antje.prasse@uniklinik-freiburg.de

Eur Respir Monogr 2013; 62: 90–95.
Copyright ERS 2013.
DOI: 10.1183/1025448x.10045012
Print ISBN: 978-1-84984-044-6
Online ISBN: 978-1-84984-045-3
Print ISSN: 1025-448x
Online ISSN: 2075-6674

The term "biomarker" is an abbreviation for "biological marker" and is used to describe parameters that can be used for health- and disease-related assessments. The National Institutes of Health (Bethesda, MD, USA) define biomarkers as characteristics that are objectively measured and evaluated as indicators of physiological processes, pathogenic processes or pharmacological responses to therapeutic intervention [1]. This chapter focuses on biological biomarkers measurable in body fluids or tissues, which are quantifiable using laboratory chemistry. While it is indisputable that other biomarkers, such as lung function test parameters and parameters derived from imaging, have a deep impact on clinical trials and daily practice in pulmonary medicine, their discussion is beyond the scope of this chapter. In the context of this chapter, clinical trials are defined as studies testing treatment-related effects in humans.

Biomarkers tested by laboratory chemistry are applied in clinical trials to test either disease- or drug-related mechanisms. There are biomarkers for disease stratification that help to delineate outcomes such as mortality and/or disease progression. Other biomarkers may indicate the risk of an individual developing a distinct event, such as acute exacerbation. While biomarkers are widely used to define patients who will benefit from a certain treatment strategy, only a few biomarkers tested by laboratory chemistry are currently implemented as surrogate markers for outcome and treatment response. Drug-related biomarkers are applied to test treatment efficacy. These biomarkers may be related to the target addressed by the tested treatment strategy (pharmacodynamics) or measure adverse reactions such as development of specific antibodies against compounds. Furthermore, many biomarkers are used for safety monitoring in clinical trials. For lung diseases, however, there is no available biomarker tested by laboratory chemistry that serves as a safety parameter.

In the last decade, several biomarkers have been implemented in the daily practice of respiratory physicians as routine laboratory parameters, and these have substantially changed decisions about treatment. It has been shown that procalcitonin is useful in the guidance of antibiotic treatment in suspected lower respiratory tract infection (LRTI), while D-dimers are a useful tool in the screening for lung embolism. Moreover, many clinicians utilise pro-brain natriuretic peptide and troponins in their diagnostic algorithm to differentiate dyspnoea induced by cardiac stress from pulmonary disease. Thus, the progress in biomarker research in the last decades has been tremendous, but advances in omics technologies will further boost this rapidly evolving field. Although most clinical trials currently test single biomarkers, we expect that omics technologies will change this to testing biomarker profiles that consist of multiple biomarkers. Testing biomarker profiles rather than single biomarkers is a further step towards developing individual treatment concepts. In the rest of this chapter, biomarkers tested in clinical trials will be discussed in detail with respect to the lung disease for which they are used.

Nonsmall cell lung cancer

Clinical trials for lung cancer took a lead in the development of biomarker-based trial protocols [2]. This was initiated by the finding that the efficacy of targeted drug treatment, such as epidermal growth factor receptor (EGFR) inhibition, is highly correlated with tumour-associated mutations in the respective target, *e.g.* EGFR [3]. Activating mutations in the tumour cell DNA of the EGFR gene confer hypersensitivity to the tyrosine kinase inhibitors gefitinib and erlotinib in patients with advanced nonsmall cell lung cancer (NSCLC) [4]. Biomarker-driven clinical trials are now largely implemented in lung cancer research. A recently published survey showed a tremendous increase in clinical trials registered at ClinicalTrials.gov that utilised biomarker-based treatment selection [5]. Most of the clinical trials that tested molecularly targeted drugs for lung cancer analysed EGFR mutations, while only a few studies addressed anaplastic lymphoma kinase (ALK) rearrangement. Moreover, novel targets such as the phosphoinositide 3-kinase (PI3K)/mammalian target of rapamycin (mTOR) pathway, the receptor tyrosine kinase MET and heat-shock proteins have been identified and respective biomarkers have been developed in parallel. In total, one quarter of all clinical trials for lung cancer used biomarker-based patient selection criteria in 2012 and one half of all clinical trials for lung cancer reported biomarker analysis [5]. For many years, tumour markers such as carcinoembryonic antigen (CEA) have been used widely in clinical trials as surrogate markers and secondary end-points to gauge treatment efficacy [6, 7]. However, positron emission tomography and computed tomography imaging is clearly superior in detecting treatment responses compared with biomarkers derived from peripheral blood.

Rare diseases

Similar to NSCLC, recent studies have reported that BRAF mutations are associated with Langerhans' cell histiocytosis [8, 9] and Erdheim–Chester disease [10] and predict treatment response to BRAF inhibitors [11]. For a Scandinavian cohort of sarcoidosis patients, BERLIN *et al.* [12] and GRUNEWALD *et al.* [13] showed that patients carrying the HLADRB1*03 gene have a favourable outcome and should not be treated with any immunosuppressive treatment, while patients with HLADRB1*15 have a high risk for persistent disease. Furthermore, it was reported recently in a small single-centre study that the response to anti-tumour necrosis factor (TNF)-α treatment may rely on the distinct promoter polymorphism G-308A in the TNF-α gene in sarcoidosis [14]. Although widely used by many clinicians, angiotensin-converting enzyme, soluble interleukin (IL)-2 receptor and neopterin levels have not been tested in clinical trials for treatment decision-making in sarcoidosis. In idiopathic pulmonary fibrosis, several biomarkers, such as the mucin MUC5b and matrix metalloproteinase MMP7 promoter polymorphisms, serum proteins and fibrocytes, have been reported to predict the outcome; however, none has been used in clinical trials so far [15–19].

LRTI

Procalcitonin is a prime example of a protein biomarker that defines patients who will benefit from a certain treatment strategy and, conversely, patients with no need for treatment [20]. Procalcitonin is a peptide precursor of the hormone calcitonin and is produced by thyroid cells and by neuroendocrine cells of the lung and the intestine. Triggered by bacterial infections, lung and intestinal neuroendocrine cells respond with an increase in their procalcitonin production, which results in elevated procalcitonin serum levels. Neither viral infections nor non-infectious inflammatory processes lead to an increase in procalcitonin levels. In 2005, the US Food and Drug Administration approved procalcitonin, in combination with other markers, as a biomarker for risk stratification in sepsis. The first study, which showed that procalcitonin measurements and guidance of antibiotic treatment reduced antibiotic prescription in 243 patients with LRTI by almost 50%, was published in 2004 [21]. The ProHOSP trial (Procalcitonin-guided Antibiotic Therapy and Hospitalisation in Patients with LRTI), including 1359 patients, confirmed these data and reported that adverse effects induced by antibiotics were significantly reduced by procalcitonin guidance [22]. Furthermore, the ProCAP study (Procalcitonin-guided Reduction of the Duration of Antibiotic Therapy in Community-acquired Pneumonia) tested community-acquired pneumonia and documented that procalcitonin-guided antibiotic duration was shortened by 65% from 12.9 to 5.8 days with a similar outcome [23]. Ventilator-associated pneumonia was also detected by procalcitonin [24, 25], which predicted survival if combined with C-reactive protein (CRP) levels [26]. In addition, antibiotic therapy exposure was significantly reduced with procalcitonin guidance in ventilator-associated pneumonia [27].

Asthma

Bronchial asthma is a disease in which tremendous efforts have been made to develop biomarkers for stratification of patients in clinical studies. Enrichment design in clinical studies is defined as a study design in which potential benefits are restricted to a subset of patients who are identified by a specific biomarker. In asthma, this approach is based on the concept of different asthma phenotypes. Bronchial asthma is a heterogeneous disease with wide variation in onset, age, sex and prognosis; therefore, asthma phenotypes describe patients with similar characteristics. In this section, we will focus on established phenotypes like allergic, eosinophilic and neutrophilic asthma [28, 29], as there are biomarker-driven clinical studies relating to each of these phenotypes. The INNOVATE study (Investigation of Omalizumab in Severe Asthma Treatment) is the most important study of omalizumab therapy in severe allergic asthma [30]. Besides a positive prick test to perennial aeroallergens, the inclusion criterion for the allergic asthma phenotype was elevated total IgE. As patients with low total IgE were not included, the INNOVATE study is an enrichment design study and a positive effect was only assumed for patients with elevated IgE. An interesting development is the use of IgE levels to monitor the omalizumab therapy. Drawbacks were the fact that IgE and omalizumab form complexes, which then lead to high total IgE values in ordinary IgE ELISAs. The development of ELISA or solid-phase immunoenzymatic assays specific for free IgE have made it possible to evaluate the treatment effect of omalizumab [31–33]. Through the measurement of free IgE, it will be possible to identify symptomatic, underdosed patients under omalizumab treatment or overdosed patients with side-effects. Considerable further research is necessary to implement free IgE for therapy monitoring in clinical practice.

Eosinophil granulocytes are also widely used in clinical trials as a biomarker for asthma. Several anti-IL-5 or anti-IL-5-receptor antibodies (mepolizumab, reslizumab and benralizumab) as well as the anti-IL-4 antibody dupilumab were especially developed for eosinophilic asthma. Multiple phase II or phase III studies are ongoing and all use elevated eosinophil granulocytes as an inclusion criterion, either in peripheral blood or in sputum [34–36]. Again, eosinophil granulocytes serve as a biomarker to allow an enrichment design study.

A further widely used biomarker in asthma is the fraction of exhaled nitric oxide (*Fe*NO). In inflammation, nitric oxide is produced by inducible nitric oxide synthases in different inflammatory cells, *e.g.* in eosinophil granulocytes. *Fe*NO measurement is conducted in patients with suspected asthma, in monitoring of disease activity and to adjust treatment with inhaled corticosteroids (ICS) [37, 38]. However, CALHOUN *et al.* [39] showed in the BASALT trial (Best Adjustment Strategy for Asthma in the Long Term) that *Fe*NO-adjusted and physician-adjusted ICS treatment are equal in time to treatment failure.

Recently, periostin was introduced as an additional biomarker for asthma that is induced by IL-13 and produced by the bronchial epithelium. Lebrikizumab, an anti-IL-13 antibody, showed promising results in the phase II trial MILLY (a study of lebrikizumab (MILR1444A) in adult patients with asthma who are inadequately controlled on ICS) [40]. This study used total IgE and peripheral blood eosinophil counts to adjust for the level of T-helper cell (Th) type 2 activation. Periostin levels were used as a surrogate marker for Th2 activity. As envisaged by the study protocol, the treatment effect of lebrikizumab was significantly more pronounced in the high periostin group. Phase III data are not available yet, but the study is designed to select for high periostin asthma patients.

Chronic obstructive pulmonary disease

Although multiple studies have investigated the role of biomarkers in chronic obstructive pulmonary disease (COPD), to date no clinical trial has used biomarkers for treatment stratification of patients with COPD. However, biomarkers show promising results predicting survival or acute exacerbations in COPD. One of the most important studies in the last few years is the ECLIPSE study (Evaluation of COPD Longitudinally to Identify Predictive Surrogate Endpoints) [41, 42]. This study identified the following biomarkers for predicting survival of patients with COPD: high-sensitivity CRP, surfactant protein D, CCL18, neutrophils, IL-6 and fibrinogen. However, C statistics showed that the added predictive power of a single biomarker to the basic clinical model of age, BODE index (body mass index, airflow obstruction, dyspnoea and exercise capacity) and hospitalisation is only minimal. In a recent study, THOMSEN *et al.* [43] confirmed that fibrinogen and leukocyte count predicted the risk for acute exacerbation in patients with COPD.

BAFADHEL *et al.* [44] conducted an elegant study, where they compared patients with acute exacerbations of COPD receiving prednisolone only, if the eosinophil granulocyte count was above 2%, *versus* standard care. In the overall analysis they showed a non-inferiority, whereas in the patients with low eosinophil counts, prednisolone treatment led to a significantly higher treatment failure and less improvement in the chronic respiratory questionnaire.

In the ProHOSP trial, SCHUETZ *et al.* [45] compared procalcitonin-based antibiotic treatment *versus* standard care in patients with LRTI. In the COPD subgroup, procalcitonin-directed antibiotic use led to decreased prescription of antibiotics, shorter duration of intravenous and oral antibiotic therapy and a non-inferiority in overall outcome. In conclusion, biomarkers seem to be useful to identify patients prone to acute exacerbation who may benefit most from treatment, but further studies are required. The predictive value of several of the above-mentioned biomarkers is significant, but clinical characteristics like age, BODE or frequent acute exacerbation or hospitalisation appear more relevant.

Conclusion

The establishment and implementation of biomarkers in clinical trials is rapidly evolving in various lung diseases. In the last decade, many biomarkers have entered the field of lung disease and some of them have already been established in daily routine. So far, most clinical trials have applied biomarkers to select for distinct disease phenotypes or the presence of the respective

treatment target. Although there is considerable interest in using biomarkers as surrogate markers for disease outcome, no pulmonary clinical trial has been published so far that has tested a biomarker as a primary end-point indicating disease progression. Omics technologies are expected to speed up discovery of and increase the number of biomarkers used in lung diseases. In the future, most pulmonary clinical trials will probably implement biomarkers.

Statement of Interest

None declared.

References

1. Foundation for the National Institutes of Health. Biomarkers Consortium. www.fnih.org/work/key-initiatives/biomarkers-consortium
2. Gazdar AF. Personalized medicine and inhibition of EGFR signaling in lung cancer. *N Engl J Med* 2009; 361: 1018–1020.
3. Toyooka S, Kiura K, Mitsudomi T. EGFR mutation and response of lung cancer to gefitinib. *N Engl J Med* 2005; 352: 2136.
4. Rosell R, Moran T, Queralt C, *et al.* Screening for epidermal growth factor receptor mutations in lung cancer. *N Engl J Med* 2009; 361: 958–967.
5. Subramanian J, Regenbogen T, Nagaraj G, *et al.* Review of ongoing clinical trials in non-small-cell lung cancer: a status report for 2012 from the ClinicalTrials.gov web site. *J Thorac Oncol* 2013; 8: 860–865.
6. Chantapet P, Riantawan P, Lebnak P, *et al.* Utility of serum cytokeratin 19 fragment (CYFRA 21-1) and carcinoembryonic antigen (CEA) as tumour markers for non-small cell lung cancer. *J Med Assoc Thai* 2000; 83: 383–391.
7. Mitsudomi T, Kosaka T, Endoh H, *et al.* Mutations of the epidermal growth factor receptor gene predict prolonged survival after gefitinib treatment in patients with non-small-cell lung cancer with postoperative recurrence. *J Clin Oncol* 2005; 23: 2513–2520.
8. Sahm F, Capper D, Preusser M, *et al.* BRAFV600E mutant protein is expressed in cells of variable maturation in Langerhans cell histiocytosis. *Blood* 2012; 120: e28–e34.
9. Yousem SA, Dacic S, Nikiforov YE, *et al.* Pulmonary Langerhans cell histiocytosis: profiling of multifocal tumors using next-generation sequencing identifies concordant occurrence of BRAF V600E mutations. *Chest* 2013; 143: 1679–1684.
10. Haroche J, Charlotte F, Arnaud L, *et al.* High prevalence of BRAF V600E mutations in Erdheim-Chester disease but not in other non-Langerhans cell histiocytoses. *Blood* 2012; 120: 2700–2703.
11. Haroche J, Cohen-Aubart F, Emile JF, *et al.* Dramatic efficacy of vemurafenib in both multisystemic and refractory Erdheim-Chester disease and Langerhans cell histiocytosis harboring the BRAF V600E mutation. *Blood* 2013; 121: 1495–1500.
12. Berlin M, Fogdell-Hahn A, Olerup O, *et al.* HLA-DR predicts the prognosis in Scandinavian patients with pulmonary sarcoidosis. *Am J Respir Crit Care Med* 1997; 156: 1601–1605.
13. Grunewald J, Eklund A, Olerup O. Human leukocyte antigen class I alleles and the disease course in sarcoidosis patients. *Am J Respir Crit Care Med* 2004; 169: 696–702.
14. Wijnen PA, Cremers JP, Erckens RP, *et al.* TNF-α G-308A polymorphism associated with response to TNF inhibitors in sarcoidosis. Paris, 6th International WASOG Conference on Diffuse Parenchymal Lung Diseases, 2013; Abstract SAR 572013.
15. Seibold MA, Wise AL, Speer MC, *et al.* A common MUC5B promoter polymorphism and pulmonary fibrosis. *N Engl J Med* 2011; 364: 1503–1512.
16. Zhang Y, Noth I, Garcia JG, *et al.* A variant in the promoter of MUC5B and idiopathic pulmonary fibrosis. *N Engl J Med* 2011; 364: 1576–1577.
17. Peljto AL, Zhang Y, Fingerlin TE, *et al.* Association between the MUC5B promoter polymorphism and survival in patients with idiopathic pulmonary fibrosis. *JAMA* 2013; 309: 2232–2239.
18. Prasse A, Probst C, Bargagli E, *et al.* Serum CC-chemokine ligand 18 concentration predicts outcome in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2009; 179: 717–723.
19. Moeller A, Gilpin SE, Ask K, *et al.* Circulating fibrocytes are an indicator of poor prognosis in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2009; 179: 588–594.
20. Schuetz P, Muller B, Christ-Crain M, *et al.* Procalcitonin to initiate or discontinue antibiotics in acute respiratory tract infections. *Cochrane Database Syst Rev* 2012; 9: CD007498.
21. Christ-Crain M, Jaccard-Stolz D, Bingisser R, *et al.* Effect of procalcitonin-guided treatment on antibiotic use and outcome in lower respiratory tract infections: cluster-randomised, single-blinded intervention trial. *Lancet* 2004; 363: 600–607.
22. Schuetz P, Christ-Crain M, Thomann R, *et al.* Effect of procalcitonin-based guidelines *vs* standard guidelines on antibiotic use in lower respiratory tract infections: the ProHOSP randomized controlled trial. *JAMA* 2009; 302: 1059–1066.

23. Christ-Crain M, Stolz D, Bingisser R, *et al.* Procalcitonin guidance of antibiotic therapy in community-acquired pneumonia: a randomized trial. *Am J Respir Crit Care Med* 2006; 174: 84–93.
24. Luyt CE, Guerin V, Combes A, *et al.* Procalcitonin kinetics as a prognostic marker of ventilator-associated pneumonia. *Am J Respir Crit Care Med* 2005; 171: 48–53.
25. Ramirez P, Garcia MA, Ferrer M, *et al.* Sequential measurements of procalcitonin levels in diagnosing ventilator-associated pneumonia. *Eur Respir J* 2008; 31: 356–362.
26. Hillas G, Vassilakopoulos T, Plantza P, *et al.* C-reactive protein and procalcitonin as predictors of survival and septic shock in ventilator-associated pneumonia. *Eur Respir J* 2010; 35: 805–811.
27. Stolz D, Smyrnios N, Eggimann P, *et al.* Procalcitonin for reduced antibiotic exposure in ventilator-associated pneumonia: a randomised study. *Eur Respir J* 2009; 34: 1364–1375.
28. Bel EH. Smoking: a neglected cause of glucocorticoid resistance in asthma. *Am J Respir Crit Care Med* 2003; 168: 1265–1266.
29. Wenzel S. Severe asthma: from characteristics to phenotypes to endotypes. *Clin Exp Allergy* 2012; 42: 650–658.
30. Humbert M, Beasley R, Ayres J, *et al.* Benefits of omalizumab as add-on therapy in patients with severe persistent asthma who are inadequately controlled despite best available therapy (GINA 2002 step 4 treatment): INNOVATE. *Allergy* 2005; 60: 309–316.
31. Klyucheva M, von Berg A, Gappa M, *et al.* Omalizumab-Therapie bei Jugendlichen mit schwerem allergischem Asthma – eine Anwendungsbeobachtung [Omalizumab therapy in adolescents with severe allergic asthma – results of a post-marketing surveillance]. *Pneumologie* 2013; 67: 233–237.
32. Korn S, Haasler I, Fliedner F, *et al.* Monitoring free serum IgE in severe asthma patients treated with omalizumab. *Respir Med* 2012; 106: 1494–1500.
33. Hamilton RG, Marcotte GV, Saini SS. Immunological methods for quantifying free and total serum IgE levels in allergy patients receiving omalizumab (Xolair) therapy. *J Immunol Methods* 2005; 303: 81–91.
34. Haldar P, Brightling CE, Hargadon B, *et al.* Mepolizumab and exacerbations of refractory eosinophilic asthma. *N Engl J Med* 2009; 360: 973–984.
35. Flood-Page P, Swenson C, Faiferman I, *et al.* A study to evaluate safety and efficacy of mepolizumab in patients with moderate persistent asthma. *Am J Respir Crit Care Med* 2007; 176: 1062–1071.
36. Castro M, Mathur S, Hargreave F, *et al.* Reslizumab for poorly controlled, eosinophilic asthma: a randomized, placebo-controlled study. *Am J Respir Crit Care Med* 2011; 184: 1125–1132.
37. Dweik RA, Boggs PB, Erzurum SC, *et al.* An official ATS clinical practice guideline: interpretation of exhaled nitric oxide levels (*FeNO*) for clinical applications. *Am J Respir Crit Care Med* 2011; 184: 602–615.
38. Barnes PJ, Dweik RA, Gelb AF, *et al.* Exhaled nitric oxide in pulmonary diseases: a comprehensive review. *Chest* 2010; 138: 682–692.
39. Calhoun WJ, Ameredes BT, King TS, *et al.* Comparison of physician-, biomarker-, and symptom-based strategies for adjustment of inhaled corticosteroid therapy in adults with asthma: the BASALT randomized controlled trial. *JAMA* 2012; 308: 987–997.
40. Corren J, Lemanske RF, Hanania NA, *et al.* Lebrikizumab treatment in adults with asthma. *N Engl J Med* 2011; 365: 1088–1098.
41. Celli BR, Locantore N, Yates J, *et al.* Inflammatory biomarkers improve clinical prediction of mortality in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2012; 185: 1065–1072.
42. Hurst JR, Vestbo J, Anzueto A, *et al.* Susceptibility to exacerbation in chronic obstructive pulmonary disease. *N Engl J Med* 2010; 363: 1128–1138.
43. Thomsen M, Ingebrigtsen TS, Marott JL, *et al.* Inflammatory biomarkers and exacerbations in chronic obstructive pulmonary disease. *JAMA* 2013; 309: 2353–2361.
44. Bafadhel M, McKenna S, Terry S, *et al.* Blood eosinophils to direct corticosteroid treatment of exacerbations of chronic obstructive pulmonary disease: a randomized placebo-controlled trial. *Am J Respir Crit Care Med* 2012; 186: 48–55.
45. Schuetz P, Suter-Widmer I, Chaudri A, *et al.* Prognostic value of procalcitonin in community-acquired pneumonia. *Eur Respir J* 2011; 37: 384–392.

Chapter 8

Health status measurement

Paul W. Jones

SUMMARY: Health status measurements are used as a marker of impaired quality of life and now form an integral part of clinical trials. They complement markers of pathophysiological processes and can provide an overall measure of the impact of symptoms and the level of benefit due to treatment. In this chapter, chronic obstructive pulmonary disease is taken as an example. A number of disease-specific measures are available, which differ in method of development, size and structure. They can provide important clinical insights and generate questions about mechanisms of benefit. The older instruments are now being recognised by regulatory authorities as valid outcome measures in clinical trials, but some of the new shorter instruments may eventually replace them. Methods of analysing health status data in the context of clinical trials are still evolving and have to deal with issues of clinical trial effects and the effect of early drop-out. There are particular challenges associated with comparison of the magnitude of difference in effect between active drugs.

Division of Clinical Science, St George's, University of London, London, UK.

Correspondence: P.W. Jones, Division of Clinical Science, St George's, University of London, London, SW17 0RE, UK. Email: pjones@sgul.ac.uk

Eur Respir Monogr 2013; 62: 96–104.
Copyright ERS 2013.
DOI: 10.1183/1025448x.10045112
Print ISBN: 978-1-84984-044-6
Online ISBN: 978-1-84984-045-3
Print ISSN: 1025-448x
Online ISSN: 2075-6674

To understand the use of health status measurement in clinical trials, it is first necessary to understand the purpose of these measurements. In any disease, arguably the two key clinical outcomes are death and poor quality of life (QoL), although clearly factors such as exacerbations, chronic disease progression and poor exercise tolerance will have a major impact on these two key outcomes and may in themselves be considered clinical outcomes in their own right. It is usually very difficult to measure clinical outcomes directly or in a standardised manner (apart from death), so surrogate measures are used, often termed "markers" [1].

In clinical trials, it is important to use very reliable measures of treatment efficacy. In this context, "reliable" has a specific meaning: that the instrument performs the same way in every setting and every patient. This is very difficult to achieve when considering QoL, because the factors that determine it will be unique to each healthy individual, so when he or she becomes ill their QoL will be affected in a manner that is unique to them. There are currently no methodologies that adequately capture this inter-personal variety, although the Schedule for the Evaluation of Individual QoL (SEIQoL) approach is one that still has potential [2, 3]. For this reason, it is better to retain the terms "QoL" for the conceptually important but unmeasurable outcome of the disease and "health status" for the measurable marker that is used to provide a surrogate estimate of the outcome. In this chapter, methods of health status measurement for chronic obstructive pulmonary disease (COPD) will be the main focus.

For COPD, health status questionnaires require much development, because the disease is complex and its symptomatic impact will vary between individuals. One of the key steps in this process is item selection, since this will determine not only the validity of the instrument, but also its reliability, *i.e.* the degree to which it performs in the same way in every setting and in every patient to whom it is administered. For this reason, the item selection process should first identify all possible relevant items, which is done by interviewing patients until no further possible items are identified. The next step is to identify, among the potentially very large number of candidate items, those that are the most appropriate and common to all patients with the disease. For example, an item that may be very important to some patients (*e.g.* playing with grandchildren) will not apply to all COPD patients, so it cannot be used. Essentially, the items that are included should form a common denominator between all patients, and as a consequence the questionnaire is stripped of items that reflect patient individuality. Thus, each individual is assessed as if he or she were a ''typical'' patient. This approach is not unique to health status measurement; the application of normal ranges for forced expiratory volume in 1 s (FEV_1) assumes that every individual is typical of the population to which they belong.

Generic health status questionnaires

There are two broad types of health status questionnaires used in COPD trials: generic instruments that can be used in any condition, and disease-specific measures. Generic instruments were used often when health status measurements were first made in clinical trials, but it was learnt quite quickly that these instruments may be insensitive to otherwise effective treatments and their use alone would give a misleading estimate of the value of the treatment. Generic questionnaires are still used in COPD trials, but these are of a specific kind, termed utility instruments. Health economists favour these questionnaires, because they have been developed to quantify health on a scale from perfect health to death and, at least in theory, can quantify states worse than death. The UK National Institute of Clinical Excellence (NICE) favours the use of the EQ-5D [4] for health economic analysis, so this instrument is now quite often used in pharmaceutical industry-sponsored trials.

Generic instruments will be discussed no further, apart from noting that they may have one specifically useful application: to quantify the impact of side-effects of treatment. A study using the long-acting β_2-agonist (LABA) salmeterol showed no clear benefit from a high dose when assessed using a disease-specific health status measure, whereas the low dose did show significant benefit [5]; a similar picture was seen with formoterol, another LABA [6]. Use of a generic instrument, the SF-36 [7], in the salmeterol trial showed that the health of patients on the higher dose was marginally worse than in those on the lower dose, across a range of health domains. This suggests that the instruments may have been detecting side-effects associated with the higher dose [5].

A major development in the field of generic measurement is the PROMIS programme (Patient Reported Outcomes Measurement Information System), which provides a suite of measurement instruments for physical, mental and social well-being. This is a very comprehensive programme that provides both short-form instruments and those that use computer adaptive testing. It is based on modern psychometric methodology using item-response theory (IRT), which will be discussed later in this chapter. This approach confers significant advantages over older instruments such as the SF-36, but the same issue of low sensitivity due to general rather than disease-specific content applies. The PROMIS website provides an excellent source of material [8].

Disease-specific questionnaires

A number of disease-specific instruments have been developed and validated for use in COPD. Most, but not all, were designed to provide an overall measure of health status impairment using a total score and most also provide a profile of domain scores. In chronological order of development, the most widely used instruments are the Chronic Respiratory Questionnaire (CRQ) [9, 10],

the St George's Respiratory Questionnaire (SGRQ) [11, 12], the Clinical COPD Questionnaire (CCQ) [13] and, most recently, the COPD Assessment Test (CAT) [14, 15]. Few clinical trial data have been published using the latter instrument, but it was developed in conformity with US Food and Drug Administration (FDA) guidance about the development and use of patient-reported outcomes in clinical trials [16] and is now being used in many pharmaceutical trials. The original CRQ was administered by an interviewer, but a self-reported version has been validated, although the two cannot be used interchangeably [17]. The SGRQ has recently been shortened specifically for use in COPD (SGRQ-C), but whilst the content was changed slightly, the scoring algorithm still produces SGRQ scores [18], so the generic title SGRQ will be used here, regardless of the version administered. The key features of each instrument are described in table 1.

One of the prime reasons for measuring health status in trials is to obtain an overall measure of the symptomatic impact of the disease. This is a methodological challenge, not least because different items may indicate different levels of severity. Indeed that is the case, as demonstrated by empirical determination of different item weights for the SGRQ [19, 20]. It will be noted that the CRQ does not have a total score and, although some researchers have combined its domain scores to produce a total score, to my knowledge, G.H. Guyatt, the developer of the instrument, has never sanctioned this approach. The CRQ and CCQ were developed using classical test theory (CTT). This is a suite of tests that, when combined, are used to provide confidence that the new questionnaire has reliable measurement properties. It should be appreciated that using CTT there is no single sufficient test; it is the summation of supportive evidence that is important. The SGRQ was also developed using CTT, but unlike the Likert scales used for the CRQ and CCQ, its items each have a unique empirically determined weight, which takes into account differences in severity of the items. The domain and total scores are calculated as a simple sum of the individual items, which goes some way towards but does not completely address the proportionate contribution of the different items to overall impaired health status.

The most recent methodology brought into questionnaire development uses IRT, the most widely used approach in medicine being Rasch modelling [21]. This method tests each item for the quality of its fit to all the other items. CTT does this in a limited way, but Rasch modelling does it more directly and over the entire scaling range of the items. This is a unique property that also allows for tests of bias due to factors such as sex and language. As part of the sequence of tests in Rasch modelling, there are also tests of the ordering of response options: essentially it tests whether the choice of response categories for each item is valid and reliable. This is rarely, if ever, done in CTT. Finally, if all the Rasch model requirements are met, a considerable degree of confidence can be applied to the conclusion that its total score is a valid overall measure and that the instrument has "invariant" measurement properties. This means, among other things, that the severity

Table 1. Key characteristics of disease-specific measures commonly used in chronic obstructive pulmonary disease (COPD) trials

	Items n	Typical completion time min	Total score	Domain scores	Response options and scoring methods
CRQ [9, 10]					
Interview	20	15–30	No	Yes	Likert scales
Self-reported	20	5–10	No	Yes	Likert scales
SGRQ [11, 12]					
Standard	50	10–15	Yes	Yes	Largely dichotomous weighted items
COPD form	40	8–10	Yes	Yes	Largely dichotomous weighted items
CCQ [13]	10	2	Yes	Yes	Likert scales
CAT [14, 15]	8	2	Yes	No	Semantic differential scale

CRQ: Chronic Respiratory Questionnaire; SGRQ: St George's Respiratory Questionnaire; CCQ: Clinical COPD Questionnaire; CAT: COPD Assessment Test.

indicated by any given distance between two points would be the same regardless of the point along the scale at which those points lie. The CAT was the first questionnaire in pulmonary medicine to use this methodology [14], but it is now being applied to questionnaires for other areas of respiratory disease [22, 23]. A significant advantage of Rasch methodology is that it allows the development of small efficient questionnaires made up of the very best core items that characterise the construct being measured. Ideally, Rasch modelling should be applied during a questionnaire's development; however, retrospective application of this methodology to the SGRQ permitted the removal of a number of less good items, which should mean that measurement properties of the SGRQ-C version should be a little better than the original [18].

The domains of those questionnaires that have them look superficially attractive, but they should be used with care. First, they are a measure of only a limited aspect of the disease; secondly, they have of necessity a smaller number of items than the whole instrument, which can lead to lower repeatability with a low signal-to-noise ratio; thirdly, it may lead to *post hoc* selection of the domain of choice to report the key health status result. On balance, it may be better to use total scores for hypothesis testing (*e.g.* "does treatment X improve health status?") and reserve the domains for *post hoc* or explanatory analysis (*e.g.* "does treatment X improve health status mainly by improving levels of activity?").

Other patient-reported outcomes

Health status measures are designed to integrate a wide range of different aspects of COPD using one score as a "black-box" measurement, so their use may be supplemented by more specific measures such as for dyspnoea. It is beyond the scope of this chapter to discuss such measures, but they have recently been reviewed elsewhere [24], and in this context it is worth noting that a new instrument has been developed to measure shortness of breath with activity specifically for use in clinical trials [25].

Health status measurements for regulatory purposes

The European Medicines Agency, in its guidance for COPD trials [26], recommends the regular use of disease-specific instruments and specifically mentions the CRQ, SGRQ and CAT, although it also says that others can be used if sufficiently validated and extensively used. The FDA in the USA has recently indicated its acceptance of the SGRQ as an outcome measure for COPD trials and this instrument is part of the the COPD Biomarker Qualification Consortium (CBQC) programme, which is a collaboration between the pharmaceutical industry, regulatory authorities and academia to identify measurement tools that can support the development of new therapies for COPD [27]. As part of that programme, a very large database of individual patient SGRQ data and other COPD markers has been assembled from patients across the world who have been recruited to a range of clinical trials and long-term observational studies. This will provide the opportunity for a very detailed analysis of many aspects of the behaviour of the SGRQ in clinical studies.

Application of health status measurement to clinical trials

As with all measurements, the application of health status measurements to a clinical trial will depend on the primary purpose of the study. While a health status score can provide an overall measure of the impact of the disease on a patient's health and well-being, it gives very little information about mechanisms. Therefore, these questionnaires need to be accompanied by markers of pathophysiological processes that can account for the mechanism of benefit, although under some circumstances it may be entirely appropriate to use a health status score as the primary outcome.

The simplest and most reported application of health status questionnaires is a pair of measurements at baseline and end of study; however, measurements at intermediate time-points also have value. From a statistical perspective, in the event of early withdrawal, an intermediate point estimate permits some form of interpolation of the missing value, at a very minimum the last number being carried forward to the end-point analysis. From a clinical perspective, multiple measurements provide useful information about the time-course of response, particularly to give an indication of when a noticeable improvement has occurred. The published literature suggests that response time (at least with the SGRQ) seems to vary between treatment types. With long-acting bronchodilators, specifically long-acting anti-muscarinic agents (LAMA), the improvement appears to continue for up to 6 months [28, 29], as shown in the example in figure 1; with inhaled corticosteroid (ICS) and LABA in combination, the response appears to be a little faster, typically 1–2 months [30, 31], as in figure 2. In one study with oral corticosteroid plus LABA, a very rapid improvement was seen over 2 weeks [32]. Therefore, it appears prudent to ''front load'' studies with more health status measurements in the first 6 months.

After approximately 6 months, health status scores begin to worsen [33, 34], but the treatment effect remained present for whole 3–4-year study period of these two studies. This is an important clinical observation, because after 2–3 years of active treatment the patient's health status may have returned to its pre-treatment state, but that is not evidence of loss of treatment efficacy; instead it indicates the presence of disease progression.

Health status measurements may also provide data that generate or provoke mechanistic hypotheses. For example, in a study to compare the effect of LAMA with ICS+LABA on exacerbations (the primary outcome of the trial) [35], there was no difference between the two treatment arms, but the health status score showed a small consistent benefit associated with ICS+LABA treatment compared with the LAMA and there was also a small but significant mortality benefit with ICS+LABA. In another double-blind study designed to test non-inferiority between a LABA and LAMA on FEV1 [36], the effects on the primary outcome were almost identical: both the SGRQ and the Transition Dyspnoea Index showed significantly greater improvement with the LABA. In both of these examples, the health status questionnaire appears to have responded to treatment effects due to mechanisms that were not linked to the trial's primary outcome.

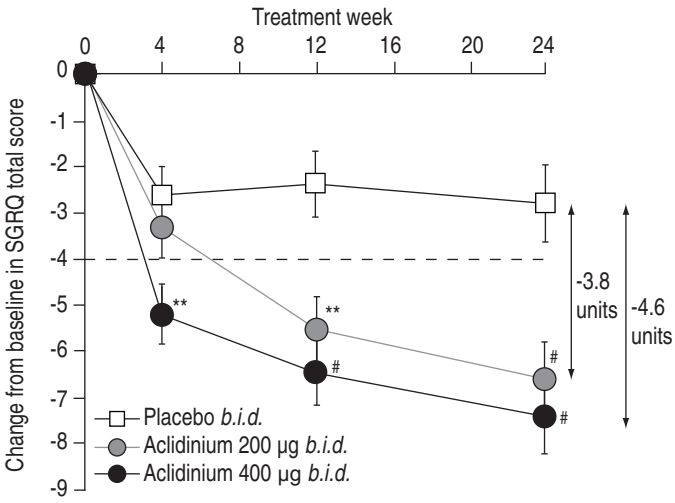


Figure 1. Change in St George's Respiratory Questionnaire (SGRQ) score with two doses of aclidinium. Data are presented as least squares mean±SE. Values on or below the dotted line represent clinically significant improvement. *b.i.d.*: twice daily. **: $p<0.01$; #: $p<0.0001$ *versus* placebo. Reproduced from [29].

Health status changes with placebo in clinical trials and study drop-outs

An improvement in health status score is consistently seen in patients in the placebo limb of clinical trials and this may continue for up to 6 months [29, 30]. Typically it is about 2–3 units on the SGRQ, which is approximately half the minimum clinically important difference (MCID), but can be larger. This is probably multifactorial in origin and unlikely to be simply due to a placebo response when answering the questionnaire. There may be a true Hawthorne effect, in which the patient changes their health behaviour because they are being observed. For example, they may

become more consistent in their use of permitted concomitant medication. In some lower-income countries, these effects may be even larger [37], as shown in figure 3, possibly because on entering a trial the patients move into a healthcare system that is more similar to that in a developed economy, so their health improves.

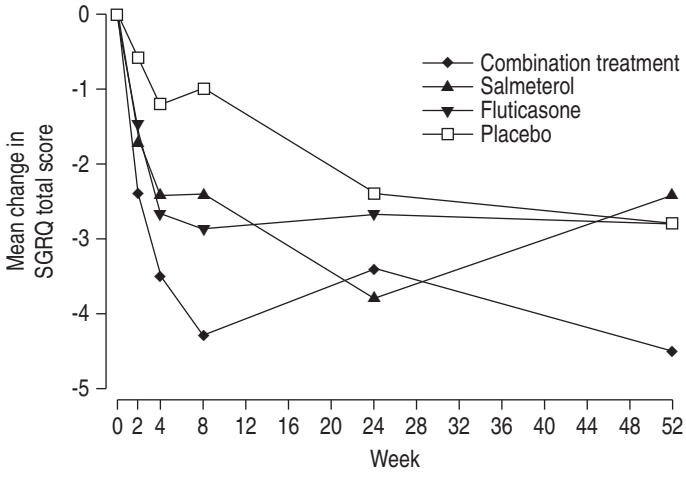


Figure 2. The effect of treatment with placebo, salmeterol, fluticasone and the combination of salmeterol and fluticasone on St George's Respiratory Questionnaire (SGRQ) score. A fall in score represents an improvement in health status. Note that the minimum important difference for the SGRQ is 4 units. Reproduced from [30] with permission from the publisher.

Studies into the changes in health status during long-term clinical trials have shown a relationship between SGRQ score and drop-out status (fig. 4) [35, 37]. It appears that patients are more likely to withdraw early if they have worse health status at entry to the study and/or their health status deteriorates whilst they are in the trial. These observations have important implications for the estimation of the size of treatment effects since, if this pattern occurs in both limbs of a study and is coupled with a greater number of drop-outs with the less effective treatment, the patients remaining in that arm will constitute a group of "healthy survivors" with better health status. This situation may result in a conservative estimate of the treatment effect.

Using MCIDs

MCID values are available for most of the health status questionnaires used in clinical trials. Issues around the 4-unit MCID for the SGRQ have been discussed in some depth elsewhere [38, 39]. MCIDs may be used in two ways to provide an indication of the size of a treatment effect. The first is to compare population means and compare these with the MCID. A drug is judged to produce a minimally important improvement if the mean change (compared with placebo) exceeds the MCID. If the data are normally distributed this means that $\geqslant$50% of treated patients must improve by an amount equal to or greater than the MCID, compared to patients treated with placebo. This is a significant treatment challenge for a condition such as COPD, so it is perhaps unsurprising that until recently the mean improvement with pharmacological therapy has usually been less than the MCID. However, recent studies have shown

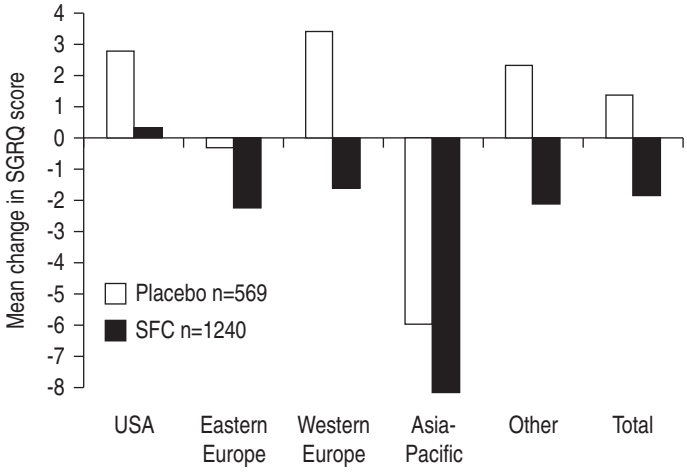


Figure 3. Mean change in St George's Respiratory Questionnaire (SGRQ) total score at 3 years by region for patients treated by placebo and salmeterol+fluticasone combination (SFC). A lower score indicates better health. A test for an interaction between region and treatment effect was not significant (p=0.16). Reproduced from [37] with permission from the publisher.

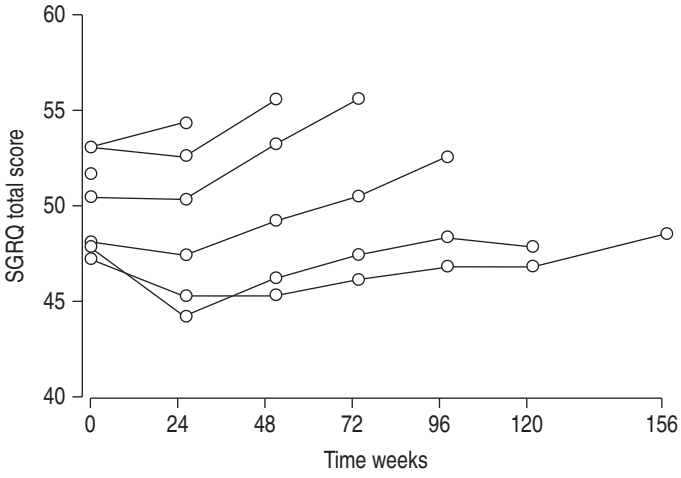


Figure 4. St George's Respiratory Questionnaire (SGRQ) scores in the placebo limb of the TORCH study (Towards a Revolution in COPD Health) in patients categorised by duration in the study. Note that only 55% of patients in this limb completed the study. Reproduced from [37] with permission from the publisher.

that mean improvements greater than the MCID are possible with long-acting bronchodilators [29, 40] and large improvements (>1.5 times the MCID) are seen with the CRQ and SGRQ in rehabilitation studies [41].

An alternative approach to comparing means, and one that may have more bearing on clinical practice, is to perform a responder analysis, by calculating the proportion of patients who exceeded the MCID. This has the advantage that the choice of score that indicates the MCID has relatively little effect on the proportion of patients who respond to active treatment compared to those who respond to placebo, over a reasonable range of MCID values [38, 42]. The disadvantage is that, because of the clinical trial effects seen in the placebo arm of studies, a large proportion of patients on placebo (typically 30–40%) have an improvement in SGRQ that exceeds the MCID, which is then to be contrasted with the 50–55% who may exceed the MCID with active treatment.

Use of MCIDs provides physicians and regulatory authorities with an index of the magnitude of health status benefit with a treatment (at least in the context of the trial). Until recently, most studies in COPD have compared active dug with placebo or short-acting bronchodilator; however, treatment in this disease is moving towards combination therapy using multiple agents, so more studies will be comparing two active drugs. It is likely that the size of difference between the drug under test and an active comparator is going to be smaller than has been seen between active drug and placebo. As a result, the difference in mean health status score will fall some way below the MCID. This has already led to questioning of MCID values or calls for lower thresholds to be used for comparisons between active drugs. However, such calls reflect a misunderstanding of the nature of the MCID, which will remain the same regardless of the therapies being compared. The smaller size of effect in comparisons between active drugs reflects the fact that existing treatments, such long-acting bronchodilators, are effective and adding a second agent of the same class is likely to come close to the maximum physiological improvement possible on that specific mechanism (*e.g.* smooth muscle relaxation). This situation should encourage a further move away from comparing health status population means to a responder analysis approach, in which the additional proportion of patients exceeding the MCID with a combination therapy, as opposed to single-agent treatment, is estimated. That of course will not remove the problem, but will shift the discussion to a debate of what constitutes a worthwhile increase in proportion of responders.

Statement of Interest

P.W. Jones reports grants and personal fees from GlaxoSmithKline, and personal fees from AZ, Novartis, UBC and Almirall, all paid to his institution.

References

1. Jones PW, Agusti AG. Outcomes and markers in the assessment of chronic obstructive pulmonary disease. *Eur Respir J* 2006; 27: 822–832.
2. O'Boyle CA, McGee H, Hickey A, *et al.* Individual quality of life in patients undergoing hip replacement. *Lancet* 1992; 339: 1088–1091.

3. Hickey AM, Bury G, O'Boyle CA, *et al.* A new short form individual quality of life measure (SEIQoL-DW): application in a cohort of individuals with HIV/AIDS. *BMJ* 1996; 313: 29–33.
4. EuroQol. EQ-5D, 2013. www.euroqol.org
5. Jones PW, Bosh TK. Quality of life changes in COPD patients treated with salmeterol. *Am J Respir Crit Care Med* 1997; 155: 1283–1289.
6. Dahl R, Greefhorst LA, Nowak D, *et al.* Inhaled formoterol dry powder *versus* ipratropium bromide in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2001; 164: 778–784.
7. Ware JE Jr, Gandek B. Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project. *J Clin Epidemiol* 1998; 51: 903–912.
8. PROMIS: Dynamic tools to measure health outcomes from the patient perspective, 2013. www.nihpromis.org
9. Guyatt GH, Berman LB, Pugsley SO, *et al.* Development of a responsive measure of quality of life for patients with chronic cardiorespiratory disease. *Clin Res* 1984; 32: 222.
10. Guyatt GH, Berman LB, Townsend M, *et al.* A measure of quality of life for clinical trials in chronic lung disease. *Thorax* 1987; 42: 773–778.
11. Jones PW, Quirk FH, Baveystock CM. The St George's Respiratory Questionnaire. *Respir Med* 1991; 85: Suppl. B, 25–31.
12. Jones PW, Quirk FH, Baveystock CM, *et al.* A self-complete measure of health status for chronic airflow limitation. The St George's Respiratory Questionnaire. *Am Rev Respir Dis* 1992; 145: 1321–1327.
13. van der Molen T, Willemse BW, Schokker S, *et al.* Development, validity and responsiveness of the Clinical COPD Questionnaire. *Health Qual Life Outcomes* 2003; 1: 13.
14. Jones PW, Harding G, Berry P, *et al.* Development and first validation of the COPD Assessment Test. *Eur Respir J* 2009; 34: 648–654.
15. Jones PW, Brusselle G, Dal Negro RW, *et al.* Properties of the COPD assessment test in a cross-sectional European study. *Eur Respir J* 2011; 38: 29–35.
16. US Dept of Health and Human Services Food and Drug Administraion, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research, *et al.* Guidance for Industry. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims, 2009. Available from: www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf
17. Williams JE, Singh SJ, Sewell L, *et al.* Development of a self-reported Chronic Respiratory Questionnaire (CRQ-SR). *Thorax* 2001; 56: 954–959.
18. Meguro M, Barley EA, Spencer S, *et al.* Development and validation of an improved COPD-specific version of the St George's Respiratory Questionnare. *Chest* 2007; 132: 456–463.
19. Quirk FH, Jones PW. Patients' perception of distress due to symptoms and effects of asthma on daily living and an investigation of possible influential factors. *Clin Sci (Lond)* 1990; 79: 17–21.
20. Quirk FH, Baveystock CM, Wilson R, *et al.* Influence of demographic and disease related factors on the degree of distress associated with symptoms and restrictions on daily living due to asthma in six countries. *Eur Respir J* 1991; 4: 167–171.
21. Andrich D. Rasch Models for Measurement. London, Sage Publications, 1988.
22. Yorke J, Moosavi SH, Shuldham C, *et al.* Quantification of dyspnoea using descriptors: development and initial testing of the Dyspnoea-12. *Thorax* 2010; 65: 21–26.
23. Patel AS, Siegert RJ, Brignall K, *et al.* The development and validation of the King's Brief Interstitial Lung Disease (K-BILD) health status questionnaire. *Thorax* 2012; 67: 804–810.
24. Jones P, Miravitlles M, van der Molen T, *et al.* Beyond FEV1 in COPD: a review of patient-reported outcomes and their measurement. *Int J Chron Obstruct Pulmon Dis* 2012; 7: 697–709.
25. Howard K, Berry P, Petrillo J, *et al.* Development of the Shortness of Breath with Daily Activities questionnaire (SOBDA). *Value Health* 2012; 15: 1042–1050.
26. European Medicines Agency. Guideline on clinical investigation of medicinal products in the treatment of chronic obstructive pulmonary disease (COPD), 2012. Available from: www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/08/WC500130880.pdf
27. Casaburi R, Celli B, Crapo J, *et al.* The COPD Biomarker Qualification Consortium (CBQC). *COPD* 2013; 10: 367–377.
28. Vincken W, van Noord JA, Greefhorst AP, *et al.* Improved health outcomes in patients with COPD during 1 yr's treatment with tiotropium. *Eur Respir J* 2002; 19: 209–216.
29. Jones PW, Singh D, Bateman ED, *et al.* Efficacy and safety of twice-daily aclidinium bromide in COPD patients: the ATTAIN study. *Eur Respir J* 2012; 40: 830–836.
30. Calverley PM, Pauwels R, Vestbo J, *et al.* Combined salmeterol and fluticasone in the treatment of chronic obstructive pulmonary disease: a randomised controlled trial. *Lancet* 2003; 361: 449–456.
31. Zheng JP, Yang L, Wu YM, *et al.* The efficacy and safety of combination salmeterol (50 μg)/fluticasone propionate (500 μg) inhalation twice daily *via* accuhaler in Chinese patients with COPD. *Chest* 2007; 132: 1756–1763.
32. Calverley PM, Boonsawat W, Cseke Z, *et al.* Maintenance therapy with budesonide and formoterol in chronic obstructive pulmonary disease. *Eur Respir J* 2003; 22: 912–919.
33. Calverley PM, Anderson JA, Celli B, *et al.* Salmeterol and fluticasone propionate and survival in chronic obstructive pulmonary disease. *N Engl J Med* 2007; 356: 775–789.

34. Tashkin DP, Celli B, Senn S, *et al.* A 4-year trial of tiotropium in chronic obstructive pulmonary disease. *N Engl J Med* 2008; 359: 1543–1554.
35. Wedzicha JA, Calverley PM, Seemungal TA, *et al.* The prevention of chronic obstructive pulmonary disease exacerbations by salmeterol/fluticasone propionate or tiotropium bromide. *Am J Respir Crit Care Med* 2008; 177: 19–26.
36. Buhl R, Dunn LJ, Disdier C, *et al.* Blinded 12-week comparison of once-daily indacaterol and tiotropium in COPD. *Eur Respir J* 2011; 38: 797–803.
37. Jones PW, Anderson JA, Calverley PM, *et al.* Health status in the TORCH study of COPD: treatment efficacy and other determinants of change. *Respir Res* 2011; 12: 71.
38. Jones PW. Interpreting thresholds for a clinically significant change in health status in asthma and COPD. *Eur Respir J* 2002; 19: 398–404.
39. Jones PW. St George's Respiratory Questionnaire: MCID. *COPD* 2005; 2: 75–79.
40. Dahl R, Chung KF, Buhl R, *et al.* Efficacy of a new once-daily long-acting inhaled β_2-agonist indacaterol *versus* twice-daily formoterol in COPD. *Thorax* 2010; 65: 473–479.
41. Lacasse Y, Goldstein R, Lasserson TJ, *et al.* Pulmonary rehabilitation for chronic obstructive pulmonary disease. *Cochrane Database Syst Rev* 2006; 4: CD003793.
42. Norman GR, Sridhar FG, Guyatt GH, *et al.* Relation of distribution- and anchor-based approaches in interpretation of changes in health-related quality of life. *Med Care* 2001; 39: 1039–1047.

Chapter 9

Electronic patient-reported outcomes

Sonya Ludmilla Eremenco

Evidera, Inc., Bethesda, MD, USA.

Correspondence: S.L. Eremenco, Evidera, Inc., 7101 Wisconsin Avenue, Suite 600, Bethesda, MD, 20814, USA.
Email: sonya.eremenco@evidera.com

SUMMARY: This chapter provides an overview of the use of electronic patient-reported outcomes (ePRO) in clinical trials focusing on respiratory and pulmonary contexts. Benefits of ePRO data collection include less subject burden, avoidance of secondary data entry errors, easier implementation of skip patterns, date and time stamping, more accurate and complete data, and less missing data. Patient-reported outcome end-points collected in clinical trials include end-points used across therapeutic areas, such as symptoms and impact, as well as those specific to the respiratory/pulmonary context, such as rescue medication use, healthcare resource utilisation, and physiological end-points. Common modes of data collection used in clinical trials and considerations for mode selection and migration from paper to electronic modes are discussed. Finally, study design considerations when implementing ePRO in clinical trials include the frequency and timing of assessments, compliance strategies for clinical trials, and compliance monitoring. Appropriate mode selection, migration, and clinical trial implementation are key steps to realise the benefits of ePRO in clinical trials.

Eur Respir Monogr 2013; 62: 105–116.
Copyright ERS 2013.
DOI: 10.1183/1025448x.10045212
Print ISBN: 978-1-84984-044-6
Online ISBN: 978-1-84984-045-3
Print ISSN: 1025-448x
Online ISSN: 2075-6674

This chapter provides an overview of the use of electronic patient-reported outcomes (ePRO) in clinical trials with a specific focus on respiratory and pulmonary contexts. It begins with definitions of both patient-reported outcomes (PRO), as well as ePRO, and discusses the benefits of ePRO data collection in clinical trials. Next, the types of PRO end-points that are collected in clinical trials are discussed including end-points used across therapeutic areas, such as symptoms and impact, as well as those that are specific to the respiratory/pulmonary context, such as rescue medication use, healthcare resource utilisation and physiological end-points. This section also includes an overview of several respiratory and pulmonary measures that have been developed as electronic instruments from the start and are currently being implemented in clinical trials. The chapter then provides an overview of the modes of data collection used in clinical trials, and considerations for mode selection and migration from paper to electronic modes. The final section of this chapter discusses study design considerations for ePRO, specifically the frequency and timing of field-based assessments, compliance strategies used in clinical trials, and compliance monitoring.

PRO and ePRO

Definitions

The 2009 US Food and Drug Administration (FDA) Guidance for Industry, titled Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims (''PRO Guidance''), describes how the FDA reviews and evaluates PRO instruments used to support claims for medical product labelling. According to the guidance, PRO instruments can be used to capture treatment benefit or risk. This chapter will focus on the capture of treatment benefit, which is defined by the guidance as ''The effect of treatment on how a patient survives, feels, or functions'' [1]. The guidance defines a PRO as ''any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else'' [1]. The guidance goes on to recommend that ''Use of a PRO instrument is advised when measuring a concept best known by the patient or best measured from the patient perspective'' [1]. Similarly, the European Medicines Agency (EMA) Reflection Paper on the Regulatory Guidance for the Use of Health Related Quality Of Life (HRQL) Measures in the Evaluation of Medicinal Products defines a PRO as ''any outcome evaluated directly by the patient himself and based on patient's perception of a disease and its treatment(s)'' [2]. These definitions and recommendations regarding PRO instruments form the basis of how PRO end-points are used in clinical trials.

An ePRO is a PRO instrument for which the data are collected using electronic means. The guidance addresses the use of ePRO for data collection and provides recommendations regarding modifications between paper and electronic formats, as well as appropriate implementation of ePRO instruments in clinical trials. When referring to ePRO instruments, the PRO guidance makes a distinction between PRO instrument administration modes and data collection methods. According to the PRO guidance, administration mode refers to self- *versus* interviewer-administered PRO measurement, while data collection method refers to the tool used for capturing the data, such as paper-based questionnaires, web-based data entry, interactive voice response systems or any of the other ePRO devices [1]. However, in the larger PRO field, it has become customary to refer to electronic data collection by the term mode and, therefore, for simplicity in this chapter, the term mode of data collection will be used to refer to collection of ePRO instruments *via* electronic technologies.

Benefits of ePRO

The emergence of ePRO data collection has occurred due to a number of advantages of electronic data collection over paper. The advantages of using electronic data collection include less subject burden, avoidance of secondary data entry errors, easier implementation of skip patterns, date and time stamping, more accurate and complete data, and less missing data [3–11]. Researchers have also noted serious drawbacks with the use of paper to collect daily diaries: they have documented non-compliance with the data collection schedule in the protocol, including cases of both back-filling (completing diary entries for previous days, not on the actual diary day) as well as forward-filling (completing diary entries for days which had not yet occurred) [7, 12–15]. It was due to recognition of these concerns with protocol compliance and the inability to know when paper diaries are completed that the FDA guidance states ''If a patient diary or some other form of unsupervised data entry is used, we plan to review the clinical trial protocol to determine what steps are taken to ensure that patients make entries according to the clinical trial design and not, for example, just before a clinic visit when their reports will be collected'' [1]. While ePRO is not explicitly stated as the solution to this concern in the guidance, ePRO has become one of the best approaches to ensure that patients follow the clinical trial protocol in terms of when they record their responses in unsupervised settings and has been viewed favourably by the FDA.

Types of PRO end-points collected electronically in respiratory clinical trials

The types of PRO end-points collected in clinical trials, regardless of the mode of data collection, are a reflection of treatment benefit of the medical product that is intended to be demonstrated by the clinical trial. As the FDA guidance states, ''The concepts measured by PRO instruments that are most often used in support of labeling claims refer to a patient's symptoms, signs, or an aspect of functioning directly related to disease status. PRO measures often represent the effect of disease (*e.g.*, heart failure or asthma) on health and functioning from the patient perspective'' [1]. Health-related quality of life (HRQoL) has been defined as ''a patient's subjective perception of the effects of the disease and treatment(s) on daily life; well-being; and psychological, physical, and social functioning'' [16] and can be considered a PRO because it is reported directly by the patient; however, it is very challenging to use HRQoL for a label claim due to its multidimensional nature [1], so it will not be included in this section of the chapter. See chapter 8 on Health status management for more information on the use of this type of measurement tool as an end-point in clinical trials [17]. In light of the concerns regarding paper data collection, as well as the numerous advantages of electronic data collection outlined above, it is very common to collect primary and key secondary end-points in clinical trials electronically to ensure the quality and reliability of these important end-points.

It is important to note that PRO end-points used in clinical trials to support labelling claims must demonstrate content validity, as well as sound psychometric properties, to ensure they can adequately measure the concepts of interest [1]. The processes to document content validity of PRO measures have been published extensively and involve patient input in the process of concept elicitation and concept confirmation *via* qualitative research, such as focus groups and one-on-one interviews, as well as documentation of saturation of the concepts in question [1, 16, 18, 19]. It is essential that these procedures be followed to document the development of measures to assess symptom and impact end-points in clinical trials. It is also important to note that the FDA has recommended that symptoms and impact be measured separately in clinical trials so that the treatment benefit can be clearly assessed for each type of end-point [20]. For this reason, symptoms and impact will be discussed in separate sections in this chapter.

Symptoms

The FDA guidance defines a symptom in the context of a PRO instrument as ''Any subjective evidence of a disease, health condition, or treatment-related effect that can be noticed and known only by the patient'' [1]. It distinguishes symptoms from signs which are ''Any objective evidence of a disease, health condition, or treatment-related effect. Signs are usually observed and interpreted by the clinician but may be noticed and reported by the patient'' [1]. Because symptoms are noticed and known only by the patient, they are the most common end-point assessed in PRO instruments developed for labelling claims. Aspects of symptoms which are frequently measured as clinical trial end-points include severity, frequency, duration or intensity. Because symptoms may fluctuate and vary frequently, depending on the condition in question, it is also very common to assess symptoms frequently, at least once a day or in some cases multiple times a day. For example, asthma symptoms are often assessed twice a day, once in the morning, to assess night-time awakenings and symptoms, and once in the evening, to assess daytime symptoms [21, 22]. Symptoms are often collected in a diary format, which is described as a set of fewer questions or concepts than a questionnaire which are completed at least daily [23]. Because such frequent assessments typically occur outside of the clinic setting, such as at the patient's home or in another unsupervised location, using ePRO to collect symptom diaries has become commonplace in clinical trials. Paper diary cards have also been used successfully to capture symptoms; for example the London cohort study of chronic obstructive pulmonary disease (COPD) patients used paper diary cards to assess COPD symptoms and reported 86% compliance

with daily completion over a 2.5 year study of 101 patients [24]. Symptoms measured daily are often averaged over time to generate the symptom score that becomes the clinical trial end-point.

Impacts

Although not explicitly defined in the FDA guidance, impact can be considered the impact of symptoms on patients' lives (*e.g.* physical functioning and participation in day-to-day activities), including a patient's ability to perform usual activities and ability to perform various roles in their life. The measurement of function often uses the term impact to measure the impact of the condition in question on the patient's ability to function in everyday life. Measures of impact often include difficulty in performing activities or limitation in activity level to show the impact of the disease or condition on functional status. Because such activities may not occur on a daily basis, it is often necessary to assess impact less frequently to allow the opportunity for performing these activities to arise. For example, an impact scale may be assessed on a weekly basis so that the opportunity for activity occurrence and, therefore, assessment of impact on ability to perform these activities is done in a meaningful way. In some conditions, such as asthma and COPD, it may be very difficult to separate out symptoms and function into separate PRO measures because the relationship between symptoms and function is very proximal. PRO instruments assessing function or impact can easily be administered using ePRO technology, which can be programmed to present the appropriate questions using the appropriate schedule which will likely vary from a symptom diary being administered. See chapter 10 on Physical activity [25] and chapter 13 on Exercise tests [26] for more information on how these two aspects of function are measured as end-points in clinical trials.

Rescue medication use

While the assessment of symptoms and function occurs in nearly all therapeutic areas, the collection of rescue medication use is only relevant to some therapeutic areas and conditions and is not addressed in the FDA guidance. However, use of rescue medication is critical in clinical trials of respiratory conditions, such as asthma and COPD, and is routinely recorded as part of a daily diary in clinical trials [22]. These diaries routinely ask whether rescue medication was used at all and in cases where it was used, how much of the medication was used. The use of ePRO enables branching logic to be used for such questions, such that patients who answer that they did not use rescue medication during the recall period in question are not then asked to report how much was used because the second question is no longer relevant. Asking only the relevant questions reduces patient burden by shortening the duration of the assessment and improves the quality of the data by preventing conflicting responses, where a patient may say no rescue medication was taken but still report an amount if the second question is presented. The ePRO technology is also programmed with edit checks, which limit the responses that can be entered, so patients may only be able to enter two digits for a number and may be asked to check their response if an excessively high amount is reported. In addition, the ePRO technology can alert the patient if they exceed a safety threshold of a certain amount of rescue medication per day and inform them that they should contact their study investigator immediately. If the ePRO system includes automatic data transmission after the diary entries are made, it will also alert the site staff that the patient may be in danger of very severe symptoms that may require immediate attention.

Healthcare resource utilisation

Healthcare resource utilisation questions may be included in a daily diary, collected electronically, to assess whether the patient sought medical attention outside of scheduled clinic visits due to worsening symptoms or changes in treatment. These types of questions may not be a primary or key secondary end-point for label claims, but provide data regarding the economic impact of the medical product that may be necessary for regulatory submission in certain countries. In addition, for some conditions, like COPD [27], use of these additional resources may signal the occurrence

of an exacerbation or the possibility of one occurring in the near future if symptoms continue to worsen without relief from the rescue medication. Collecting this information electronically using a system with automatic data transmission can enable the clinical site to monitor the patient's status and use of such resources, so that the site can contact the patient immediately for follow-up.

Physiological end-points

There is growing interest in applying the electronic data collection and transmission capabilities of ePRO technology to the collection of physiological end-points, such as lung function in clinical trials. While another chapter in this *European Respiratory Monograph* addresses lung function as an end-point in clinical trials [28], in this section of this chapter the use of technology to collect these physiological data will be addressed. Examples of physiological data that can be collected electronically include blood glucose readings using a glucometer, peak expiratory flow (PEF) using a peak flow meter and forced expiratory volume in 1 s (FEV_1) using a spirometer. The use of peak flow meters to monitor a patient's PEF during a clinical trial is a very common end-point in asthma clinical trials, and in some COPD trials. There is a body of evidence showing that, whether these PEF readings are collected using manual or electronic peak flow meters, if the patients manually record the entries into a diary at home, there is a substantial amount of inaccuracy in these manual recordings, involving cases of errors but also intentional falsification to obscure severe readings [29, 30], and this is improved if the data are recorded electronically [31]. Therefore, in light of this evidence of the high degree of inaccuracy in manual entry of physiological data by patients, it is critical that such data are collected electronically in clinical trials, to ensure the accuracy of the data used for analysis in the trial. The section of this chapter on modes of data collection will provide further detail regarding the options for collecting and transmitting physiological data, such as PEF readings, using ePRO technology in clinical trials.

Respiratory measures developed electronically

In recent years, a number of new PRO instruments for respiratory and pulmonary clinical trials have been developed, using the principles of the FDA guidance mentioned above, which have also been developed from the start as ePRO instruments. One of the earliest such instruments was the EXAcerbations of Chronic obstructive pulmonary disease Tool (EXACT), which was developed, by a consortium of industry, academia, PRO experts and the FDA, to provide a tool designed to standardise the method for evaluating the frequency, severity and duration of acute exacerbations (AE) of COPD and chronic bronchitis in international studies and clinical trials [32–35]. In addition to using a novel approach to instrument development through consortia, the EXACT-PRO Initiative was forward thinking in the intention to develop this as an electronic daily diary administered using a handheld device, such as a smartphone or personal digital assistant (PDA). The EXACT-Respiratory Symptoms (EXACT-RS) is a separate outcome measure for use in stable COPD patients, consisting of 11 questions from EXACT that capture changes in specific respiratory symptoms (breathlessness, cough and sputum, and chest symptom domains) [35, 36], and is now being used in clinical trials to assess respiratory symptoms of stable COPD patients [35, 37]. Other recently developed ePRO instruments for COPD include the Night-time Symptoms of COPD Instrument (NiSCI) and Early Morning Symptoms of COPD (EMSCI) [38–40]. The development of ePRO instruments for asthma clinical trials is also underway, with the Critical Path Institute's PRO Consortium convening a working group to develop an asthma symptom assessment [41]. Another newly developed ePRO instrument is the Pulmonary Arterial Hypertension – Symptoms and Impact (PAH-SYMPACT), which is the first PRO instrument for pulmonary arterial hypertension, that uses a tablet electronic device for daily diary administration [42, 43]. The PAH-SYMPACT consists of a daily diary section as well as an impact section that is completed every 7 days, in line with the recommendations for assessing symptoms and impact in separate domains. These instruments demonstrate the growing trend for developing new respiratory and pulmonary instruments as ePRO instruments from the start, which has streamlined the instrument development process and improved data collection of these critical end-points in clinical trials.

Modes of data collection

A number of modes of data collection are available for clinical trial use, which include paper-and-pen as well as digital pen, various electronic devices including handheld devices, tablet devices, laptop or desktop computers and software-based technologies, such as interactive voice response (IVR), web-based systems and mobile ''apps''. Table 1 presents an overview of commonly used technologies for the collection of PRO data in clinical trials.

The one similarity amongst all the electronic modes described in table 1 is that all the data collected electronically are transmitted or entered, in the case of IVR and Web-based instruments, into a central database, which is maintained by the ePRO vendor for the duration of the study.

Integration with physiological devices

In respiratory clinical trials, an additional device type may be used to collect physiological data from patients, most commonly the peak flow meter to capture PEF readings. Peak flow meters can be integrated with ePRO devices in one of two ways. Originally, peak flow meters were separate electronic devices that could be linked to an ePRO handheld device using infrared radio frequency or Bluetooth transmission, such that the data from the peak flow meter would be transferred to the ePRO device once the connection between the devices was made, and then the PEF readings would be transmitted to a central database for storage and later retrieval for analysis. In recent years, single integrated devices which contain the peak flow meter apparatus as well as a small screen have been developed to simplify the data collection process for these two data streams, so that the same device collects both the PRO data as well as the physiological data. The benefit of the single device is the convenience of needing to use only one device and not having to worry about enabling the two devices to connect with each other. However, the screen sizes may be smaller so that the capabilities of these single devices may be more limited than today's smartphones, so it is challenging to fit diary questions and responses onto these devices and compromises may need to be made. These devices are most commonly used in asthma, where the PEF readings are an important end-point and are often collected twice daily.

Figure 1 shows examples of the types of devices that integrate the diary functionality with peak flow or spirometry measurements.

Mode selection

With such an array of ePRO technologies available, it is important to select the appropriate mode for the clinical trial in a careful way. Mode selection should take place as early as possible, after the PRO instrument of interest has been chosen for the trial. If an electronic mode is selected, there is a systematic development process that must be followed to ensure that the system is compliant with 21 Code of Federal Regulations (CFR) part 11 regulations and EMA requirements [1, 45–52], which may take 2–3 months to be completed.

The following is a list of issues to consider when selecting the appropriate mode for a given clinical trial or clinical trial programme. Frist, the characteristics of the patient population: consider whether the mode is suitable for the patient population in terms of age, physical limitations, or other limitations such as visual, hearing or cognitive impairment. Secondly, the characteristics of the instrument: the length of the instrument and the format of the responses may determine which of several suitable modes represents the best choice. For example, a long instrument that uses a visual analog scale with no numerical representation of the responses on the scale would not be suitable for IVR where the responses must be entered numerically, so a visual mode would be a better choice. Thirdly, it is necessary to consider the frequency of data collection: an episodic diary that is event-driven, meaning that episodes are to be entered immediately after they occur, will require the most portable option to make it easiest for the patient to comply with the assessment schedule. Less frequent assessment will allow for more options. Fourthly, the location of data

Table 1. Modes of data collection for patient-reported outcome (PRO) end-points in clinical trials

Mode	Description
Paper-and-pen	Oldest mode for PRO data collection. PRO instruments are printed on paper, which also requires shipping to sites in study binders and manual data entry into a database. Paper does not restrict how the patient answers and has no mechanisms to prevent skipping questions or missing data altogether. No compliance reminders available.
Digital pen	Special pen containing a camera records coordinates written on specially printed paper containing a unique dot matrix pattern. Date and time stamp of when the writing occurred is available. Similar to paper in terms of limitations around skipped responses, missing data or lack of compliance reminders.
Handheld device	Formerly PDAs, now smartphones are used to collect diary data in clinical trials. The device is self-contained with software residing on the device and can be used offline to respond to the diary and then transmit the data separately. Small screens allow one question to be displayed per screen so patient burden may increase with longer questionnaires. Devices are programmed with alarms, to remind the patient to complete the diary at a specific time, and edit checks to prevent erroneous data or skipped questions. Very commonly used for PRO data collection outside of the clinic setting because of convenience and portability, as many can fit into the patient's pocket or handbag.
Tablet device	Larger screen than a handheld device, can accommodate a larger font size and more text on the screen, this provides the option of single or multiple questions per screen. Less portable than a handheld device, but more portable than a desktop or laptop computer. Programmed in a similar way to handheld devices in that the software is self-contained on the device and the questions are completed offline and then transmitted separately. Smaller size tablets, like the iPad (Apple Distribution International, Hollyhill, Republic of Ireland) or Galaxy Tab (Samsung Electronics UK, Chertsey, UK), are becoming more attractive for diary use at the patient's home if the patient does not need to carry the device around with them. Tablet devices can be programmed with alarms and reminders if they are used outside the clinic.
Desktop/laptop device	Early ePRO studies used desktop or laptop computers to collect PRO data at clinic sites and this is still possible in some cases. These devices are not as mobile and, therefore, are most often used in a clinic setting or by patients in their own homes to complete web-based instruments.
IVR	The PRO questions and responses are audio-recorded and then read to the patient who responds by entering a number on the numeric keypad of a telephone while listening to the questions and responses. IVR systems can be programmed to call the patient to remind them to complete the diary, but this is subject to each country's privacy laws and ethics rules regarding the patient's own phone number. IVR uses a very different cognitive process of auditory rather than visual processing, which may increase the cognitive load on the patient [44].
Web-based	This is a software-based technology which allows the patient to access the PRO instrument using a web-browser and a wide range of devices which can access the web, ranging from mobile phones with internet capabilities, to tablets, laptops and desktop computers. Because of the wide range and screen size of devices which can access web-based PROs, it is critical that the screens be designed with a certain device or screen size in mind to reduce inconsistencies in how the PRO instrument is displayed across patients. Email reminders are typically used with web-based PRO instruments, but rely on the patient being online to receive the reminder and respond to the PRO in a timely fashion.
Mobile ''apps''	A recent development in the ePRO field is the development of ''apps'' or applications which can be downloaded from the Web to an individual patient's own smartphone. After download, it can be configured to run as a standalone application on the device, operating like a handheld device. There are challenges associated with ensuring consistency in how the PRO instrument is displayed across a wide range of device types, so PRO ''apps'' are not yet in wide use for regulatory clinical trials.

IVR: interactive voice response; PDA: personal digital assistant; ePRO: electronic PRO.

Figure 1. a) Asthma Monitor 1 (AM1; eResearchTechnology (ERT), Philadelphia, PA, USA) (stand-alone device). b) In2itive Model 2120 (Vitalograph Ltd., Buckingham, UK) (integrated device).

collection: whether the data are collected at clinic visits or in the field is often a major factor in the choice of mode. An associated consideration is whether there is a need for immediate access to the data: if patients need to be monitored between study visits, the ePRO technology needs to support this through data transmission in the field. Fifthly, are there any geographic issues need to be addressed: is the chosen mode feasible for use in all the countries planned for the clinical trial? Is there the appropriate infrastructure to support this mode, especially transmission *via* cellular phone signals? And finally, there is the issue of cost. While this is not a scientific issue, the modes vary widely in cost as there are software costs in all cases, but hardware costs in addition for the device-based modes.

These mode selection criteria need to be considered in an iterative way, as they are not necessarily sequential in nature but must be considered simultaneously.

Migration considerations

This section will briefly review the considerations for migration between modes, especially when one mode has been previously validated and a different mode is used in future clinical trials. The FDA guidance states that the properties of a PRO instrument are format and mode-specific and that the FDA will want to see evidence of the adequacy of the instrument in the new mode to ensure that the measurement properties still hold [1]. In response to this expectation, the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) ePRO task force published a report on assessment of measurement equivalence between modes to provide recommendations for best practices in this area [53]. The task force report stated that the degree of modification during the migration from one mode to another would determine the level of evidence needed to support measurement equivalence. Minor modifications would be supported by qualitative evidence including cognitive interviews and usability testing, moderate modifications would require equivalence testing and usability testing to ensure equivalence, and substantial modifications would be treated as a new instrument and would require full psychometric testing to demonstrate the measurement properties of the instrument [53]. There is some evidence of equivalence between paper and visual modes of data collection [54], but further empirical work is needed to document equivalence between paper and IVR modes. Another consideration when moving from paper to electronic modes is how to carry out a faithful migration, which means that patients will interpret and respond to the questions in the new mode in the same way as the old mode regardless of changes in format that may have occurred during the migration (ISPOR ePRO Task Force unpublished report on mixed modes of data collection, 2013). Finally, ensuring that the technology is easy for all patients to use is another important consideration when migrating to electronic modes. For screen-based platforms, it is necessary to ensure that the presentation is

legible, that the size of the area to be tapped is sufficient and that the navigation is clear, while for IVR the considerations have to do with voice type and speed [55].

Study design considerations

Clinical trial study designs can incorporate the capabilities of ePRO technology to facilitate better and more accurate assessment of the clinical trial end-points outlined above. These study designs need to take into consideration the nature of the condition when determining when and how the assessments should be completed, especially those that are completed outside the clinic setting.

Frequency and timing of assessments

The FDA guidance discusses frequency of assessments within a section on study design considerations and states the following: "The frequency of PRO assessment should correspond with the specific research questions being addressed, length of recall asked by the instrument's response options, demonstrated instrument measurement properties, the disease or condition's natural history, the treatment's nature, and planned data analysis. Some diseases, conditions, or clinical trial designs may necessitate more than one baseline assessment and several PRO assessments during treatment" [1].The assessment frequency can be programmed into the ePRO system so that the assessments occur at the correct time per protocol and help to ensure compliance with the protocol on the part of the patients. The visit schedule can also be programmed into the site-based ePRO system if used, which will help site staff to ensure patients are completing the correct assessments at the right time and in the right order.

Compliance strategies

The clinical trial study design also needs to include consideration of compliance strategies that can be used to enhance the patient's compliance with the ePRO assessments. Devices can be programmed with multiple alarms or reminders that are scheduled to sound at a preset time every day to help the patient maintain the appropriate completion schedule. Other compliance strategies include reinforcement, such as displaying the patient's compliance percentage over time to encourage them to continue if compliance is high or to remind them if compliance is low [56]. Patient training is also linked to compliance, because training improves the patient's ability to use the technology and reduces cases of missing an entry due to a problem like forgetting the PIN code and not being able to access the device to answer questions [57, 58]. It is also common practice to include a minimum level of compliance with a daily diary as an inclusion criterion in a clinical trial to ensure that the patient will meet a minimum standard of compliance during screening and provide sufficient data for baseline data, under the assumption that their compliance level will be similar during the clinical trial itself.

Compliance monitoring

In addition to compliance features of the ePRO system, a compliance monitoring plan needs to be included in the study design to ensure that site staff and/or the sponsor's or contract research organisation's monitor are reviewing compliance reports and taking action if compliance with ePRO data collection declines during the study. The ePRO system can be programmed to alert the site staff if a patient has not completed a PRO or transmitted data within a set number of days, so that the staff can intervene to find the source of the problem and encourage the patient to become compliant with their PRO assessment. It is also possible to create a report that allows the sponsor to see whether a site has viewed the compliance reports in order to monitor how well the site is reviewing and tracking their patients' compliance rates in the study. Both compliance strategies and compliance monitoring are important aspects of clinical trial planning to optimise patient compliance with ePRO data collection in order to yield as much high quality data as possible in the trial.

Conclusion

This chapter provided an overview of the use of ePROs in clinical trials with a specific focus on respiratory and pulmonary contexts. Definitions of both PROs as well as ePROs were provided and the benefits of ePRO data collection in clinical trials were outlined. Next, the types of PRO end-points that are collected in clinical trials were discussed including end-points used across therapeutic areas, such as symptoms and impact, as well as those that are specific to the respiratory/pulmonary context, such as rescue medication use, healthcare resource utilisation, and physiological end-points. This section also included an overview of several respiratory and pulmonary measures that have been developed as electronic instruments from the start and are currently being implemented in clinical trials. The chapter then provided an overview of modes of data collection that are used in clinical trials, along with considerations for mode selection and migration from paper to electronic modes. The final section of this chapter discussed study design considerations for use of ePRO, specifically the frequency and timing of assessments, compliance strategies used in clinical trials and compliance monitoring to ensure high quality data are collected in clinical trials.

Statement of Interest

S.L. Eremenco reports grants from Amgen, Janssen, AbbVie, Merck, Actelion, Exelixis and Shire, and grants and non-financial support from Forest, outside the submitted work.

References

1. Food and Drug Administration. Guidance for industry on patient-reported outcome measures: use in medical product development to support labeling claims. *Federal Register* 2009; 74: 65132–65133.
2. European Medicines Agency. Reflection paper on the regulatory guidance for the use of health related quality of life (HRQL) measures in the evaluation of medicinal products. London, European Medicines Agency, 2005.
3. Bloom DE. Technology, experimentation, and the quality of survey data. *Science* 1998; 280: 847–848.
4. Bushnell DM, Reilly MC, Galani C, *et al.* Validation of electronic data capture of the Irritable Bowel Syndrome – Quality of Life Measure, the Work Productivity and Activity Impairment Questionnaire for Irritable Bowel Syndrome and the EuroQol. *Value Health* 2006; 9: 98–105.
5. Hyland ME, Kenyon CA, Allen R, *et al.* Diary keeping in asthma: comparison of written and electronic methods. *BMJ* 1993; 306: 487–489.
6. Ross J, Ross E, Holzbaur E. Electronic patient reported outcomes (ePRO): the best defense in preventing missing PRO data. *Value Health* 2013; 16: A53.
7. Stone AA, Shiffman S, Schwartz JE, *et al.* Patient non-compliance with paper diaries. *BMJ* 2002; 324: 1193–1194.
8. Taenzer PA, Speca M, Atkinson MJ, *et al.* Computerized quality-of-life screening in an oncology clinic. *Cancer Pract* 1997; 5: 168–175.
9. Tiplady B, Crompton GK, Dewar MH, *et al.* The use of electronic diaries in respiratory studies. *Drug Inform J* 1997; 31: 759–764.
10. Tourangeau R, Smith TW. Asking sensitive questions: the impact of data collection mode, question format, and question context. *Public Opin Q* 1996; 60: 275–304.
11. Velikova G, Wright EP, Smith AB, *et al.* Automated collection of quality-of-life data: a comparison of paper and computer touch-screen questionnaires. *J Clin Oncol* 1999; 17: 998–1007.
12. Broderick JE, Schwartz JE, Shiffman S, *et al.* Signaling does not adequately improve diary compliance. *Ann Behav Med* 2003; 26: 139–148.
13. Jamison RN, Raymond SA, Levine JG, *et al.* Electronic diaries for monitoring chronic pain: 1-year validation study. *Pain* 2001; 91: 277–285.
14. Palermo TM, Valenzuela D, Stork PP. A randomized trial of electronic *versus* paper pain diaries in children: impact on compliance, accuracy, and acceptability. *Pain* 2004; 107: 213–219.
15. Stone AA, Shiffman S, Schwartz JE, *et al.* Patient compliance with paper and electronic diaries. *Control Clin Trials* 2003; 24: 182–199.
16. Patrick DL, Burke LB, Gwaltney CJ, *et al.* Content validity – establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 1 – eliciting concepts for a new PRO instrument. *Value Health* 2011; 14: 967–977.
17. Jones PW. Health status measurement. *In:* Kolb M, Vogelmeier CF, eds. Outcomes in Clinical Trials. *Eur Respir Monogr* 2013; 62: 96–104.

18. Leidy NK, Vernon M. Perspectives on patient-reported outcomes: content validity and qualitative research in a changing clinical trial environment. *Pharmacoeconomics* 2008; 26: 363–370.
19. Patrick DL, Burke LB, Gwaltney CJ, *et al.* Content validity – establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2 – assessing respondent understanding. *Value Health* 2011; 14: 978–988.
20. Food and Drug Administration. Measurement in Clinical Trials: Review and Qualification of Clinical Outcome Assessments; Public Workshop October 19, 2011. Silver Spring, Food and Drug Administration, 2011.
21. Santanello NC, Barber BL, Reiss TF, *et al.* Measurement characteristics of two asthma symptom diary scales for use in clinical trials. *Eur Respir J* 1997; 10: 646–651.
22. Svedsater H, Clark M, Martin S, *et al.* Measurement properties of an asthma symptom and rescue medication use diary: a critical review. *Am J Respir Crit Care Med* 2013; 187: A4213.
23. Raymond SA. Forward – the patient's viewpoint: impact of ePRO technologies on clinical research. *In:* Byrom B, Tiplady B, eds. ePro: Electronic Solutions for Patient-Reported Data. Farnham, Gower Publishing Ltd, 2010; pp. xi–xxx.
24. Seemungal TA, Donaldson GC, Bhowmik A, *et al.* Time course and recovery of exacerbations in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2000; 161: 1608–1613.
25. Watz H. Physical activity. *In:* Kolb M, Vogelmeier CF, eds. Outcomes in Clinical Trials. *Eur Respir Monogr* 2013; 62: 117–126.
26. Troosters T. Exercise tests. *In:* Kolb M, Vogelmeier CF, eds. Outcomes in Clinical Trials. *Eur Respir Monogr* 2013; 62: 170–178.
27. Jones PW, Chuecos F, Lamarca R, *et al.* Unreported exacerbations of chronic obstructive pulmonary disease are associated with a reduction in health status: results from the ATTAIN study. *Am J Respir Crit Care Med* 2013; 187: A6072.
28. O'Donnell DE, Laveneziana P, Neder JA. Lung function. *In:* Kolb M, Vogelmeier CF, eds. Outcomes in Clinical Trials. *Eur Respir Monogr* 2013; 62: 127–145.
29. Kamps AW, Roorda RJ, Brand PL. Peak flow diaries in childhood asthma are unreliable. *Thorax* 2001; 56: 180–182.
30. Verschelden P, Cartier A, L'Archevêque J, *et al.* Compliance with and accuracy of daily self-assessment of peak expiratory flows (PEF) in asthmatic subjects over a three month period. *Eur Respir J* 1996; 9: 880–885.
31. Reddel HK, Toelle BG, Marks GB, *et al.* Analysis of adherence to peak flow monitoring when recording of data is electronic. *BMJ* 2002; 324: 146–147.
32. Jones PW, Chen WH, Wilcox TK, *et al.* Characterizing and quantifying the symptomatic features of COPD exacerbations. *Chest* 2011; 139: 1388–1394.
33. Leidy NK, Wilcox TK, Jones PW, *et al.* Development of the EXAcerbations of Chronic Obstructive Pulmonary Disease Tool (EXACT): a patient-reported outcome (PRO) measure. *Value Health* 2010; 13: 965–975.
34. Leidy NK, Wilcox TK, Jones PW, *et al.* Standardizing measurement of chronic obstructive pulmonary disease exacerbations. Reliability and validity of a patient-reported diary. *Am J Respir Crit Care Med* 2011; 183: 323–329.
35. Leidy NK, Murray LT. Patient-reported outcome (PRO) measures for clinical trials of COPD: the EXACT and E-RS. *COPD* 2013; 10: 393–398.
36. Sexton CC, Leidy NK, Notte SM, *et al.* Quantifying the severity of respiratory symptoms of chronic obstructive pulmonary disease (COPD): reliability and validity of a patient diary. *Am J Respir Crit Care Med* 2011; 183: A1481.
37. Beier J, Kirsten AM, Mroz R, *et al.* Improvements in COPD symptoms and rescue medication use with aclidinium bromide compared with tiotropium and placebo: a phase IIIb study. *Am J Respir Crit Care Med* 2013; 187: A4276.
38. Eremenco S, Palsgrove A, Mocarski M, *et al.* Evaluating the usability of an ePRO diary for measuring COPD symptoms at night and early morning. *Value Health* 2012; 15: A73.
39. Palsgrove A, Houghton K, Hareendran A, *et al.* The development of the early morning symptoms of COPD instrument (EMSCI). *Value Health* 2011; 14: A496–A497.
40. Palsgrove A, Schaefer M, Hareendran A, *et al.* The development of a patient-reported outcome instrument to evaluate nighttime symptoms of COPD. *Value Health* 2011; 14: A141–A142.
41. Critical Path Institute PRO. Consortium. Fourth Annual Patient-Reported Outcome (PRO) Consortium Workshop - Public Workshop, Silver Spring, MD April 24th and April 25th, 2013. http://c-path.org/wp-content/uploads/2013/09/2013_Asthma_WG_Poster.pdf Date last updated: 2013. Date last accessed: May 31, 2013.
42. McCollister D, Kummer S, Badesch D, *et al.* Development of the pulmonary arterial hypertension-symptoms and impact (PAH-SYMPACT) questionnaire: a new disease-specific patient-reported outcome instrument for PAH. *Am J Crit Care Med* 2013; 187: A3294.
43. McCollister D, Kummer S, Badesch DB, *et al.* Development of a new patient-reported outcome (PRO) instrument for pulmonary arterial hypertension (PAH): the pulmonary arterial hypertension–symptoms and impact (PAH-SYMPACT) questionnaire. *Value Health* 2013; 16: A34.
44. Sweller J. Cognitive load during problem solving: effects on learning. *Cognitive Sci* 1988; 12: 257–285.
45. European Medicines Agency (EMEA). GCP Inspectors Working Group. Reflection paper on expectations for electronic source documents used in clinical trials 2007. www.emea.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/10/WC500004385.pdf Date last updated: Oct 17, 2007. Date last accessed: May 31, 2013.
46. Food and Drug Administration. 21 CFR part 11, electronic records; electronic signatures; final rule. *Federal Register* 1997: 62: 13429–13466.

47. Food and Drug Administration. General principles of software validation; final guidance for industry and FDA staff. www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm085371.pdf Date last updated: Jan 11, 2002. Date last accessed: May 31, 2013.
48. Food and Drug Administration. Guidance for industry, Part 11, electronic records; electronic signatures, maintenance of electronic records 2002. www.21cfrpart11.com/files/fda_docs/draft_guidance_for_ind_july2002.pdf Date last updated: July 2002. Date last accessed: May 31, 2013.
49. Food and Drug Administration. Guidance for industry, Part 11, electronic records; electronic signatures – scope and application 2003. www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm072322.pdf Date last updated: August 2003. Date last accessed: May 31, 2013.
50. Food and Drug Administration. Guidance for industry. Computerized systems used in clinical trials 1999. www.fda.gov/downloads/ICECI/EnforcementActions/BioresearchMonitoring/UCM133749.pdf Date last updated: April 1999. Date last accessed: May 31, 2013.
51. Eremenco S., Revicki DA. Regulation and compliance: scientific and technical regulatory issues associated with electronic capture of patient-reported outcome data. *In:* Byrom B, Tiplady B, eds. ePro: Electronic Solutions for Patient-Reported Data. Farnham, Gower Publishing Ltd, 2010; pp. 79–106.
52. European Commission Health and Consumers Directorate-General. Eudralex. Volume 4 Good Manufacturing Practice. Medicinal Products for Human and Veterinary Use. Annex 11: Computerised Systems. http://ec.europa.eu/health/files/eudralex/vol-4/annex11_01-2011_en.pdf Date last updated: 2010. Date last accessed: May 31, 2013.
53. Coons SJ, Gwaltney CJ, Hays RD, *et al.* Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. *Value Health* 2009; 12: 419–429.
54. Gwaltney CJ, Shields AL, Shiffman S. Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. *Value Health* 2008; 11: 322–333.
55. Tiplady B. Diary design considerations: interface issues and patient acceptability. *In:* Byrom B, Tiplady B, eds. ePro: Electronic Solutions for Patient-Reported Data. Farnham, Gower Publishing Ltd, 2010; pp. 167–184.
56. Cates C, Roberts L, Eremenco S, *et al.* Can older, chronically ill adults use electronic diaries? Compliance rates in a prospective study of patients with COPD. *Value Health* 2009; 12: A126.
57. Ross J, O'Gorman H. Addressing the patient factor in ePRO to improve compliance. *J Patient Compliance* 2013; 1: 34–37.
58. Ross J, Ross E, O'Gorman HJ, *et al.* Take time to train: evidence to show that patient training in electronic patient reported outcomes is beneficial. *Value Health* 2011; 14: A151.

Chapter 10

Physical activity

Henrik Watz

SUMMARY: Many health benefits related to primary and secondary prevention of chronic diseases can be ascribed to regular physical activity. Respiratory diseases often lead to physical inactivity, which is an indicator of poor prognosis in chronic obstructive pulmonary disease. Sustained improvement of physical inactivity is a key component in the treatment of respiratory diseases and is believed to be a clinically relevant end-point. However, more research is needed to accurately measure levels of physical activity and to define clinically meaningful changes for patients. The greatest impact of regular physical activity on mortality reduction can be achieved by mild-to-moderate physical activity making physical inactivity a modifiable risk factor for non-communicable diseases, with an impact or mortality rate comparable to that of smoking.

Pulmonary Research Institute at Lung Clinic Grosshansdorf, Airway Research Center North, Member of the German Center for Lung Research, Grosshansdorf, Germany.

Correspondence: H. Watz, Pulmonary Research Institute at Lung Clinic Grosshansdorf, Airway Research Center North, Member of the German Center for Lung Research, Woehrendamm 80, D-22927 Grosshansdorf, Germany. Email: h.watz@pulmoresearch.de

Eur Respir Monogr 2013; 62: 117–126.
Copyright ERS 2013.
DOI: 10.1183/1025448x.10045312
Print ISBN: 978-1-84984-044-6
Online ISBN: 978-1-84984-045-3
Print ISSN: 1025-448x
Online ISSN: 2075-6674

Physical activity is defined as "any bodily movement produced by skeletal muscles that results in energy expenditure" [1]. Physical activity needs to be distinguished from exercise and physical fitness. Exercise is a subcategory of physical activity. It is planned, structured, repetitive and purposeful in the sense that improvement or maintenance of one or more components of physical fitness is an objective [1]. Components of physical fitness are the attributes that relate to the ability to perform physical activity: cardiorespiratory endurance, muscular endurance, muscular strength, body composition and flexibility [1]. Therefore, measurements of exercise capacity and physical fitness can only indicate what a person is capable of doing, whereas a measurement of physical activity reflects what a person actually does: domestic work, occupational activities, and leisure-time activities including sports [1, 2]. Physical activity is the sum of all these activities [1].

Physical activity also needs to be distinguished from "activities of daily living". This term refers to a set of basic everyday tasks related to personal self-care and independent living [3, 4].

Measurement of physical activity

Physical activity can be measured by: direct observation, indirect calorimetry or the doubly labelled water technique (measures of energy expenditure), questionnaires, and physical activity monitors that measure bodily movements (accelerometers) or steps (pedometers) [5].

Direct observation and indirect calorimetry, for which patients breathe through a mouthpiece or face mask, are impractical for long-term measurement of physical activity [5, 6]. The doubly labelled water technique is a very expensive method restricted to specific research questions [5].

Questionnaires for self-reported physical activity are widely used in epidemiological research. Most of our current knowledge regarding the health benefits of regular physical activity is derived

from questionnaire-based self-reports of physical activity in large cohort studies [7, 8]. Self-reported assessment of physical activity by questionnaires is often subject to recall bias [9, 10] and in most cases does not correlate very well with objectively quantified physical activity in patients with chronic obstructive pulmonary disease (COPD) [11–14]. In general, this method does not provide an accurate estimate of free-living energy expenditure [15].

Objective monitoring of physical activity can be performed with wearable activity monitors such as pedometers or accelerometers [5]. Pedometers are small, relatively inexpensive devices worn at the hip or ankle that record the number of steps walked per day [16]. Walking is something that most people can do, and steps per day is a simple metric for assessing physical activity [17], either through public health campaigns [5] or in randomised controlled trials [16]. Results of a recent systematic review suggest that the use of pedometers is associated with significant increases in physical activity and significant decreases in body mass index and blood pressure in the general population [16]. An important predictor of an increase of physical activity in all studies performed to date is having a step goal of 10 000 steps per day [16].

A 12-week pedometer-based exercise counselling strategy effectively enhanced daily physical activity, physical fitness and health-related quality of life in COPD outpatients who did not participate in a rehabilitation programme [18].

Wearable accelerometers are designed to detect acceleration as a way to measure bodily movements [5]. Accelerometers have emerged as an important means of assessing the duration and intensity of physical activity and have served to define primary outcome measures in several observational and interventional studies [19], especially in chronic disease populations such as COPD patients [20]. Currently, several types (uni-, bi- or triaxial accelerometers) generating different output measures are available [5]. Integrated multisensor systems combine accelerometry with other sensors that capture body responses to exercise (*e.g.* heart rate or skin temperature) in an attempt to optimise physical activity assessments and to obtain reasonably valid estimates of energy expenditure [5, 10, 20]. Recent research recommendations for accelerometer use [9, 21] and validation studies for various accelerometers in COPD patients [22–24] are of utmost importance for standardising accelerometer measurements so that physical activity can become an objectively measured outcome in clinical trials.

Health benefits of regular physical activity in the general population

Epidemiological research has convincingly demonstrated that regular physical activity is related to various health benefits, resulting in reduced mortality [7]. About 9% of premature mortality worldwide (5.3 million deaths per year globally) can be attributed to physical inactivity. This makes physical inactivity a relevant and modifiable risk factor for non-communicable diseases, with an impact on mortality rate comparable to that of smoking [25, 26].

Interestingly, it seems that the greatest impact of regular physical activity on mortality reduction can be achieved by mild-to-moderate physical activity, while higher physical activity levels provide little additional benefit for reduction of all-cause mortality [27].

By evaluating health benefits in a disease-specific context, it was shown that moderately intense physical activity is protective for the development of major non-communicable diseases, such as cardiovascular disease, metabolic syndrome/diabetes mellitus, cancer, osteoporosis and depression [7]. For secondary prevention, patients with multiple chronic diseases also benefit from regular physical activity with regard to their mortality risk [7, 28].

The health benefits of regular physical activity on lung function have not been studied as extensively. One epidemiological study found that moderate to high levels of regular physical activity are associated with reduced lung function decline and COPD risk among active smokers, but not in ex-smokers or those who have never smoked [29].

Although current health recommendations advocate 30 min of moderate physical activity 5 days a week [30], recent data indicate that as little as 15 min of walking 6 days a week significantly reduces all-cause mortality by 14%, cancer mortality by 10% and mortality from cardiovascular disease by 20% [8, 31].

Health benefits of regular physical activity in patients with COPD

Few studies have evaluated the relationship between reduced physical activity and risk of hospitalisation for exacerbations of COPD [32]. All studies have demonstrated a higher risk of hospitalisation for patients identified as inactive by questionnaires or by an objective, accelerometer-based measurement [32, 33].

Three studies evaluated the relationship between physical inactivity and all-cause mortality in patients with COPD [33–35] and found a strong association between the level of physical activity and the risk of death from any cause. Patients with the lowest amount of physical activity had the highest mortality irrespective of other important confounders, *e.g.* lung function impairment or functional exercise capacity [33–35]. Interestingly, physical activity parameters, assessed by energy expenditure (total daily energy expenditure divided by resting energy expenditure) or steps per day, had the strongest predictive power for mortality in COPD compared to other established predictors (fig. 1) [35]. For every 14% decrease in physical activity level (corresponding to a decrease of ~200–250 kcal in active daily energy expenditure), the relative risk of death more than doubled (fig. 1) [35].

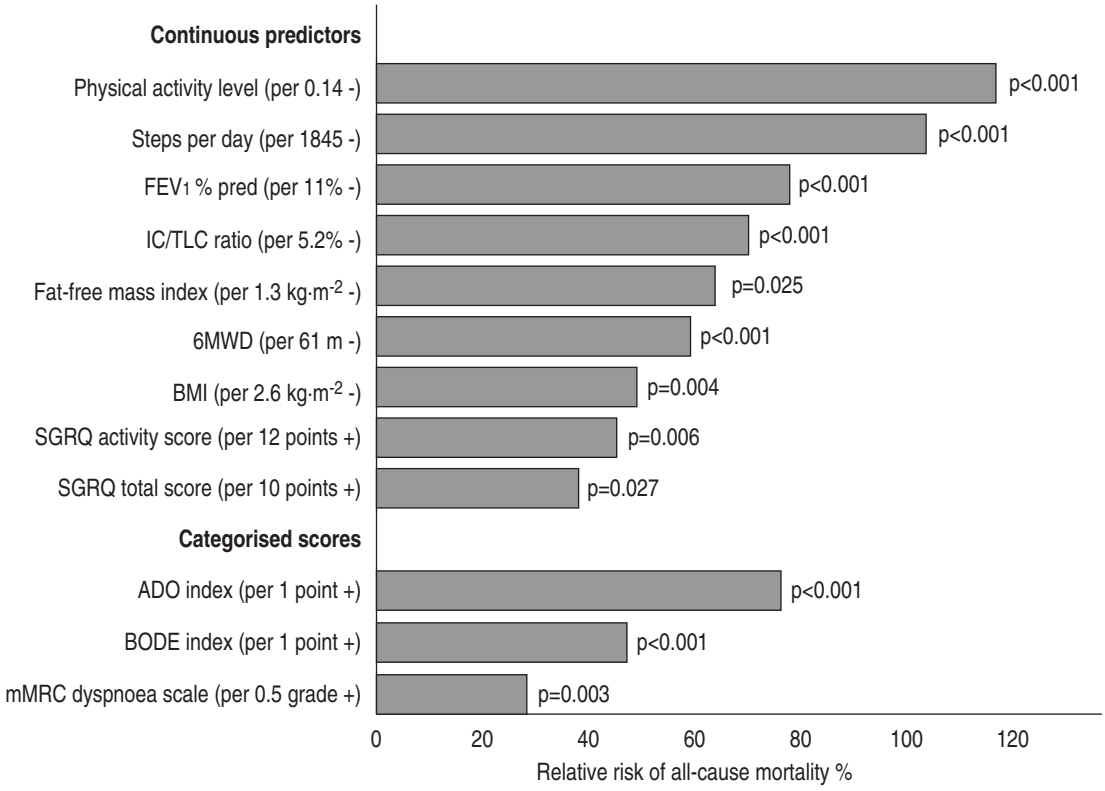


Figure 1. Relative risk of death associated with physical activity in comparison with established predictors of mortality in 170 patients. The relative risk of mortality was calculated from corresponding hazard ratios. FEV1: forced expiratory volume in 1 s; % pred: % predicted; IC: inspiratory capacity; TLC: total lung capacity; 6MWD: 6-min walking distance; BMI: body mass index; SGRQ: St George's Respiratory Questionnaire; ADO: age, dyspnoea, airflow obstruction; BODE: BMI, airway obstruction, dyspnoea, exercise capacity; mMRC: modified Medical Research Council. Reproduced from [35] with permission from the publisher.

Decrease of physical activity in patients with COPD

Objective measurements of physical activity by various activity monitors have consistently demonstrated that patients with COPD have reduced physical activity compared to healthy control subjects [36]. This holds true for the duration, intensity and counts (*e.g.* steps or movements) of physical activity parameters, although the intensity of physical activity may be the least affected [36].

A significant decrease of physical activity in patients with COPD can be observed early in the course of the disease. Two studies independently showed that physical activity, as assessed by steps per day, is reduced when beginning in Global Initiative for Chronic Obstructive Lung Disease (GOLD) stage II (fig. 2) [10, 37]. However, depending on the levels of physical activity in the control group, a significant decrease of physical activity was also observed in patients with mild COPD [38].

Physiological components related to reduced physical activity in COPD patients

Most studies reported thus far have demonstrated that the decrease in physical activity in COPD patients is related to disease severity, as indicated by the decrease in forced expiratory volume in 1 s (FEV_1) [10, 36, 39]. Generally, weak to moderate correlations were observed between FEV_1 and physical activity measures of various activity monitors [10, 11, 39–42]. Dynamic lung

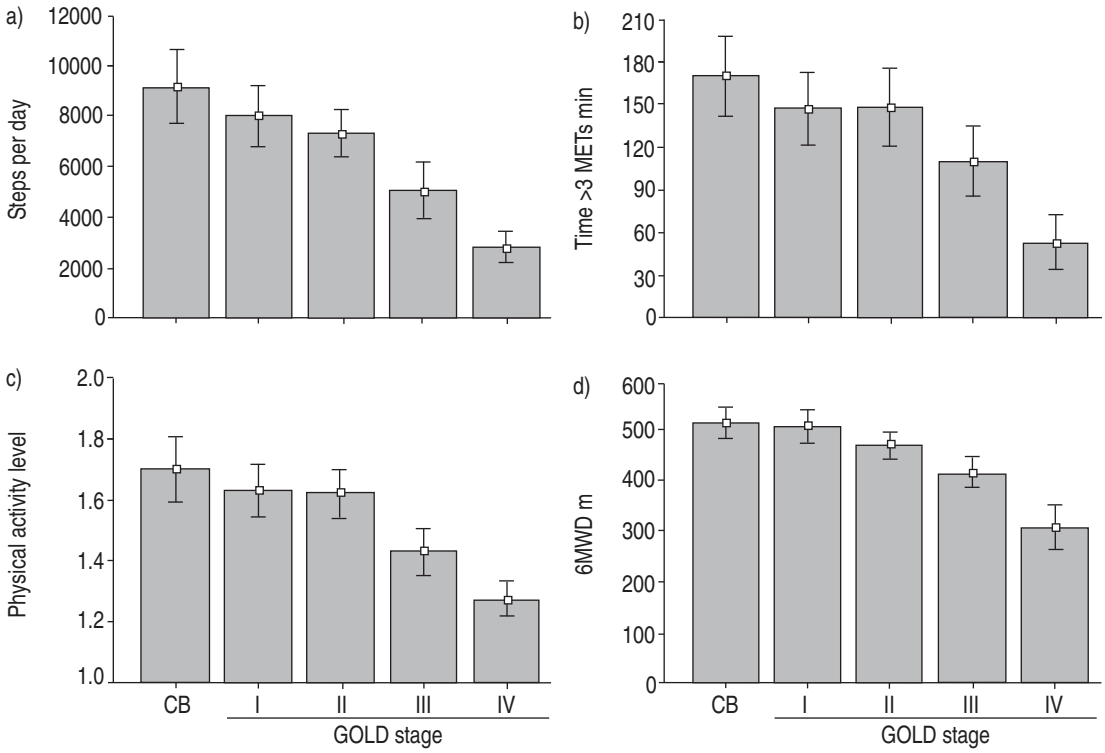


Figure 2. a) Steps per day, b) minutes >3 metabolic equivalents (METs), c) physical activity level and d) 6-min walking distance (6MWD) for patients with chronic bronchitis (CB) (former Global Initiative for Chronic Obstructive Lung Disease (GOLD) stage at risk, *i.e.* cough and sputum production in smokers or ex-smokers, normal lung function) or chronic obstructive pulmonary disease (COPD) according to GOLD stages (n=163). A significant decrease in the number of steps per day was observed at GOLD stage II for COPD patients compared with patients with CB (p=0.02). Reproduced from [10].

hyperinflation, known to be a relevant factor for exercise limitations in COPD, has been shown to be more closely related to physical activity than FEV1, static lung volumes or diffusion capacity [43].

Exercise capacity, assessed by either 6-min walking distance (6MWD) or cardiopulmonary exercise testing, is moderately related to physical activity in COPD (fig. 3) [10, 39, 41, 43]. However, the variation in physical activity between subjects at a given value of 6MWD was very large, especially when the 6MWD approached normal values (>500 m) (fig. 3) [10]. Even with a 6MWD of <350 m, indicating a clinically relevant exercise limitation, some patients still showed physical activity in the range observed in patients with normal values for the 6MWD (fig. 3). The relationship between exercise capacity and physical activity can also be seen from another perspective. Although 6MWD is only slightly reduced in GOLD stages I and II, physical activity measured with an accelerometer over a period of 5 days was significantly affected in the early stages (fig. 2). This may indicate that activity monitoring in the home could reveal health impairments due to COPD earlier than a measurement of exercise performance for 6 min at a submaximal level in the laboratory.

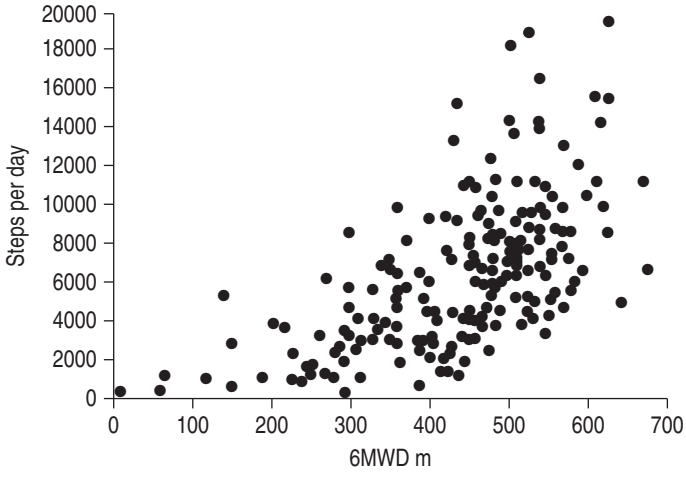


Figure 3. Relationship between 6-min walking distance (6MWD) and steps per day in patients with chronic obstructive pulmonary disease. Data from [10].

Comorbidities have also been evaluated for their impact on physical activity levels in patients with COPD. It has been shown that cardiac dysfunction, reflected by higher systemic levels of N-terminal pro-brain natriuretic peptide and echocardiographic measures of impaired left ventricular filling, is related to reduced physical activity [44]. Furthermore, patients with COPD and metabolic syndrome have less physical activity than patients with COPD without metabolic syndrome [45]. In three studies evaluating the effects of muscle dysfunction on physical activity, reduced quadriceps muscle strength was related to a decrease in physical activity [38, 39, 41]. Quadriceps wasting seems to be independently associated with physical inactivity in patients with mild COPD [38].

Low-grade systemic inflammation is related to physical inactivity in COPD patients [39, 44, 46]. This may simply reflect the known link between physical inactivity and systemic inflammation due to disuse of the muscles [47, 48].

Can physical inactivity be predicted by other characteristics of disease severity or questionnaires?

Given the relatively robust relationship between lung function or exercise capacity and physical activity in COPD, it is reasonable to speculate that lung function parameters or measures of exercise capacity may have some value for the prediction of physical inactivity in patients with COPD. However, neither FEV1 nor 6MWD, along with other established disease characteristics, were reliable predictors of physical inactivity in COPD (table 1) [10, 14]. Physical activity questionnaires were not substantially better predictors of physical inactivity than the common clinical characteristics of COPD [13].

Physical activity in respiratory diseases other than COPD

Physical activity in respiratory diseases other than COPD has not been extensively studied. One study in patients awaiting lung transplantation included patients with interstitial lung disease.

Table 1. Receiver operating characteristic curves for prediction of 163 physically very inactive[#] patients with chronic obstructive pulmonary disease (COPD)

Predictor	Area	SE	p-value	95% CI of area	Cut-off point[¶]	Sensitivity	Specificity	PPV	NPV	Likelihood ratio of test result	
										Positive	Negative
GOLD stage	0.823	0.035	<0.001	0.76–0.89	Stage III	0.80	0.75	0.69	0.85	3.25	0.26
BODE score	0.796	0.036	<0.001	0.73–0.87	3	0.67	0.79	0.69	0.78	3.23	0.42
FEV_1	0.795	0.036	<0.001	0.73–0.87	54% pred	0.80	0.71	0.65	0.84	2.78	0.28
6MWT	0.773	0.037	<0.001	0.70–0.85	445 m	0.79	0.74	0.68	0.84	3.06	0.29
mMRC	0.768	0.037	<0.001	0.70–0.84	Grade 2	0.65	0.78	0.67	0.77	3.01	0.44

PPV: positive predictive value; NPV: negative predictive value; GOLD: Global Initiative for Chronic Obstructive Lung Disease; BODE: body mass index, airway obstruction, dyspnoea, exercise capacity; FEV_1: forced expiratory volume in 1 s; 6MWT: 6-min walk test; mMRC: modified Medical Research Council dyspnoea scale; % pred: % predicted. [#]: defined as a physical activity level of <1.40, calculated by the ratio of total daily energy expenditure/resting energy expenditure (energy expenditure measured with the SenseWear armband; SMT Medical, Würzburg, Germany), including 66 patients with COPD (40.5% of all COPD patients in this study); [¶]: equal to or greater than. Data from [10].

Generally, these patients had the same level of physical inactivity as COPD patients with an FEV_1 of 25% predicted [49]. The question arises as to whether any improvements in physical activity can be achieved in patients with progressive lung disorders such as idiopathic pulmonary fibrosis. However, the recently published improvements in exercise intolerance and health-related quality of life following pulmonary rehabilitation in patients with interstitial lung disease [50] might stimulate further research into improvements in physical activity levels in these patients after an intervention.

Available data on physical activity levels in patients with asthma are inconsistent [51]. Although physical inactivity has been identified as a potential risk factor for the development of asthma [52], less is known about measured physical activity levels in established asthma, either controlled or uncontrolled. A recent survey using the short telephone version of the International Physical Activity Questionnaire did not reveal significant differences in physical activity levels between adult patients with asthma and control subjects [51]. In children, poorly controlled asthma is associated with reduced physical activity, and improved asthma control is associated with clinically relevant increases in physical activity [53].

Interventional studies to improve physical activity in COPD

Only a few studies have evaluated whether an intervention, by either pulmonary rehabilitation or pharmacological therapy, improves physical activity in patients with COPD. 10 studies assessed the effect of pulmonary rehabilitation on physical activity levels measured by activity monitors in patients with COPD [42, 54–62]. While four studies observed significant improvements in physical activity [42, 60–62], the other six studies did not show significant treatment effects on physical activity after pulmonary rehabilitation despite documented treatment effects on exercise capacity and health-related quality of life [54–59]. In one study, improvements in walking time were observed after 6 months of continuous pulmonary rehabilitation but not after 3 months. This may indicate that longer lasting programmes are needed to implement changes in behaviour [62] and that physical activity is a complex behaviour influenced by many factors within the environmental, social/cultural and psychological/cognitive domains [63, 64].

Two studies measured physical activity after long-acting bronchodilator therapy in patients with COPD

[65, 66]. In one study, significant improvements in exercise tolerance were demonstrated after therapy with the long-acting β_2-agonist indacaterol, while objectively measured physical activity did not improve when compared with placebo [65]. Another study reported significant improvements in physical activity with indacaterol treatment [66]. However, these results are difficult to interpret because treatment was open-label [66].

Physical activity compared with other established outcome parameters in clinical studies in COPD

Unlike the well-established outcome parameters applied in COPD studies that aim to objectively assess patient-oriented outcomes, the measurement of physical activity is far from being standardised, and a clear response to intervention has not been demonstrated (table 2) [67]. However, strictly speaking, improvements in exercise capacity and exertional dyspnoea are only indicative of improvement in physical activity. It is likely that health benefits of treatment can only be achieved when patients become more active in daily life [6]. As highlighted by CASABURI [6],

Table 2. Physical activity as an outcome variable: general considerations

Is the methodology standardised?
No guidelines for measurement of physical activity in respiratory disease are available
Different activity monitors and questionnaires are used
Most studies in COPD measured physical activity using an accelerometer and a measurement period of at least 2 days
Has the frequency distribution been established?
In the general population: not in every country
In COPD patients: several accelerometer studies have been conducted in various countries and settings
Are the data reproducible?
Reproducibility is not optimal, for example when: day-to-day variability has an impact on reproducibility; seasonal variation has an impact on reproducibility; and compliance/adherence with activity monitoring might have an impact on reproducibility
Has the influence on treatment been established?
No, there are conflicting results regarding the effectiveness of, for example, pulmonary rehabilitation or bronchodilator treatment on physical activity parameters
Can the measurements be performed during exacerbations of COPD?
Yes
Does the variable influence common comorbidities, *e.g.* ischaemic heart disease, diabetes mellitus or psychiatric disturbances?
Regular physical activity is effective in primary and secondary prevention of non-communicable diseases such as ischaemic heart disease, diabetes mellitus and depression
Can the variable be used in multicentre trials?
Yes
In order to measure the variable, what is needed?
A questionnaire: depending on the research question
Equipment: activity monitors such as pedometers and accelerometers enable investigators to objectively quantify physical activity
Time: measurement needs to be performed over several days, *e.g.* 1 week
Training: most activity monitors are easy to use
How much does the measurement of the variable cost?
Cheap, *e.g.* questionnaires and pedometers
Moderate, *e.g.* accelerometers
Expensive, *e.g.* some types of accelerometers
Are there any safety issues?
No
Are there any ethical issues?
No

COPD: chronic obstructive pulmonary disease. Reproduced with modification from [67].

Table 3. Practical approach to measuring physical activity in clinical trials in chronic obstructive pulmonary disease

Triaxial accelerometers
Actigraph GT3X (Actigraph, Pensacola, FL, USA) [23, 24]
Dynaport MiniMod (McRoberts BV, The Hague, the Netherlands) [23, 24]
SenseWear armband (Body Media Inc., Pittsburgh, PA, USA) [22, 23]
Measurement period
1 week, including the weekend [10]
Valid measurement (minimal wearing time of the accelerometer required for inclusion of the day in the data analyses)
80% of the predefined wearing time (*e.g.* waking day, 12 h, 18 h, or 24 h) [9, 21]

a ''paradigm shift'' is needed to promote physical activity and measurement of physical activity as part of the clinically meaningful assessments in clinical studies.

Physical activity as an outcome in COPD research: a practical approach for accelerometer measurements

Based on current data regarding the validity of accelerometers [22–24] and the reliability of the measurement period [9, 10, 21], a practical approach for measuring physical activity is presented in table 3.

What might be a minimal important difference for improvements in physical activity?

A minimal important difference for improvements in physical activity does not exist. 30 min of moderate physical activity per day is currently advocated by health recommendations [30]. It is likely that in most patients with severe respiratory disease, this goal cannot be achieved. A study that demonstrated health benefits with reduced mortality achieved by only 15 min of moderate physical activity per day [8] should give us cause for cautious optimism that clinically relevant changes can be achieved in patients with severe respiratory disease.

Statement of Interest

H. Watz has received speaking honorarium from AstraZeneca, Boehringer, Novartis, Almirall and Berlin Chemie, and is also on the advisory board for Boehringer. Outside the submitted work, H. Watz used the SenseWear armband in several studies; the Sensewear armband was provided by SMT Medical (Würzburg, Germany).

Acknowledgements

I would like to thank Mary McKenney for critical review of the revised manuscript.

References

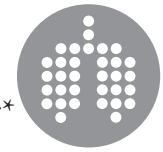
1. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Rep* 1985; 100: 126–131.
2. Hartman JE, Boezen HM, de Greef MH, *et al.* Consequences of physical inactivity in chronic obstructive pulmonary disease. *Expert Rev Respir Med* 2010; 4: 735–745.
3. Katz S. Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living. *J Am Geriatr Soc* 1983; 31: 721–727.
4. Fricke J. Activities of daily living. *In:* JH Stone, M Blouin, eds. International Encyclopedia of Rehabilitation. 2013. Available from: http://cirrie.buffalo.edu/encyclopedia/en/article/37/
5. Pitta F, Troosters T, Probst VS, *et al.* Quantifying physical activity in daily life with questionnaires and motion sensors in COPD. *Eur Respir J* 2006; 27: 1040–1055.

6. Casaburi R. Activity promotion: a paradigm shift for chronic obstructive pulmonary disease therapeutics. *Proc Am Thorac Soc* 2011; 8: 334–337.
7. Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. *CMAJ* 2006; 174: 801–809.
8. Wen CP, Wai JP, Tsai MK, *et al.* Minimum amount of physical activity for reduced mortality and extended life expectancy: a prospective cohort study. *Lancet* 2011; 378: 1244–1253.
9. Ward DS, Evenson KR, Vaughn A, *et al.* Accelerometer use in physical activity: best practices and research recommendations. *Med Sci Sports Exerc* 2005; 37: Suppl. 1, S582–S588.
10. Watz H, Waschki B, Meyer T, *et al.* Physical activity in patients with COPD. *Eur Respir J* 2009; 33: 262–272.
11. Steele BG, Holt L, Belza B, *et al.* Quantitating physical activity in COPD using a triaxial accelerometer. *Chest* 2000; 117: 1359–1367.
12. Pitta F, Troosters T, Spruit MA, *et al.* Activity monitoring for assessment of physical activities in daily life in patients with chronic obstructive pulmonary disease. *Arch Phys Med Rehabil* 2005; 86: 1979–1985.
13. Garfield BE, Canavan JL, Smith CJ, *et al.* Stanford Seven-Day Physical Activity Recall questionnaire in COPD. *Eur Respir J* 2012; 40: 356–362.
14. van Gestel AJ, Clarenbach CF, Stowhas AC, *et al.* Predicting daily physical activity in patients with chronic obstructive pulmonary disease. *PLoS One* 2012; 7: e48081.
15. Manini TM, Everhart JE, Patel KV, *et al.* Daily activity energy expenditure and mortality among older adults. *JAMA* 2006; 296: 171–179.
16. Bravata DM, Smith-Spangler C, Sundaram V, *et al.* Using pedometers to increase physical activity and improve health: a systematic review. *JAMA* 2007; 298: 2296–2304.
17. Moy ML, Danilack VA, Weston NA, *et al.* Daily step counts in a US cohort with COPD. *Respir Med* 2012; 106: 962–969.
18. Hospes G, Bossenbroek L, ten Hacken NH, *et al.* Enhancement of daily physical activity increases physical fitness of outclinic COPD patients: results of an exercise counseling program. *Patient Educ Couns* 2009; 75: 274–278.
19. Catellier DJ, Hannan PJ, Murray DM, *et al.* Imputation of missing data when measuring physical activity by accelerometry. *Med Sci Sports Exerc* 2005; 37: Suppl. 11, S555–S562.
20. Van Remoortel H, Giavedoni S, Raste Y, *et al.* Validity of activity monitors in health and chronic disease: a systematic review. *Int J Behav Nutr Phys Act* 2012; 9: 84.
21. Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. *Med Sci Sports Exerc* 2005; 37: Suppl. 11, S531–S543.
22. Hill K, Dolmage TE, Woon L, *et al.* Measurement properties of the SenseWear armband in adults with chronic obstructive pulmonary disease. *Thorax* 2010; 65: 486–491.
23. Van Remoortel H, Raste Y, Louvaris Z, *et al.* Validity of six activity monitors in chronic obstructive pulmonary disease: a comparison with indirect calorimetry. *PLoS One* 2012; 7: e39198.
24. Rabinovich RA, Louvaris Z, Raste Y, *et al.* Validity of physical activity monitors during daily life in patients with COPD. *Eur Respir J* 2013; 42: 1205–1215.
25. Lee IM, Shiroma EJ, Lobelo F, *et al.* Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet* 2012; 380: 219–229.
26. Wen CP, Wu X. Stressing harms of physical inactivity to promote exercise. *Lancet* 2012; 380: 192–193.
27. Lollgen H, Bockenhoff A, Knapp G. Physical activity and all-cause mortality: an updated meta-analysis with different intensity categories. *Int J Sports Med* 2009; 30: 213–224.
28. Martinson BC, O'Connor PJ, Pronk NP. Physical inactivity and short-term all-cause mortality in adults with chronic disease. *Arch Intern Med* 2001; 161: 1173–1180.
29. Garcia-Aymerich J, Lange P, Benet M, *et al.* Regular physical activity modifies smoking-related lung function decline and reduces risk of chronic obstructive pulmonary disease: a population-based cohort study. *Am J Respir Crit Care Med* 2007; 175: 458–463.
30. Haskell WL, Lee IM, Pate RR, *et al.* Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Circulation* 2007; 116: 1081–1093.
31. Nigam A, Juneau M. Survival benefit associated with low-level physical activity. *Lancet* 2011; 378: 1202–1203.
32. Seidel D, Cheung A, Suh ES, *et al.* Physical inactivity and risk of hospitalisation for chronic obstructive pulmonary disease. *Int J Tuberc Lung Dis* 2012; 16: 1015–1019.
33. Garcia-Rio F, Rojo B, Casitas R, *et al.* Prognostic value of the objective measurement of daily physical activity in patients with COPD. *Chest* 2012; 142: 338–346.
34. Garcia-Aymerich J, Lange P, Benet M, *et al.* Regular physical activity reduces hospital admission and mortality in chronic obstructive pulmonary disease: a population based cohort study. *Thorax* 2006; 61: 772–778.
35. Waschki B, Kirsten A, Holz O, *et al.* Physical activity is the strongest predictor of all-cause mortality in patients with COPD: a prospective cohort study. *Chest* 2011; 140: 331–342.
36. Vorrink SN, Kort HS, Troosters T, *et al.* Level of daily physical activity in individuals with COPD compared with healthy controls. *Respir Res* 2011; 12: 33.
37. Troosters T, Sciurba F, Battaglia S, *et al.* Physical inactivity in patients with COPD, a controlled multi-center pilot-study. *Respir Med* 2010; 104: 1005–1011.
38. Shrikrishna D, Patel M, Tanner RJ, *et al.* Quadriceps wasting and physical inactivity in patients with COPD. *Eur Respir J* 2012; 40: 1115–1122.

39. Waschki B, Spruit MA, Watz H, *et al.* Physical activity monitoring in COPD: compliance and associations with clinical characteristics in a multicenter study. *Respir Med* 2012; 106: 522–530.
40. Belza B, Steele BG, Hunziker J, *et al.* Correlates of physical activity in chronic obstructive pulmonary disease. *Nurs Res* 2001; 50: 195–202.
41. Pitta F, Troosters T, Spruit MA, *et al.* Characteristics of physical activities in daily life in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2005; 171: 972–977.
42. Walker PP, Burnett A, Flavahan PW, *et al.* Lower limb activity and its determinants in COPD. *Thorax* 2008; 63: 683–689.
43. Garcia-Rio F, Lores V, Mediano O, *et al.* Daily physical activity in patients with chronic obstructive pulmonary disease is mainly associated with dynamic hyperinflation. *Am J Respir Crit Care Med* 2009; 180: 506–512.
44. Watz H, Waschki B, Boehme C, *et al.* Extrapulmonary effects of chronic obstructive pulmonary disease on physical activity: a cross-sectional study. *Am J Respir Crit Care Med* 2008; 177: 743–751.
45. Watz H, Waschki B, Kirsten A, *et al.* The metabolic syndrome in patients with chronic bronchitis and COPD: frequency and associated consequences for systemic inflammation and physical inactivity. *Chest* 2009; 136: 1039–1046.
46. Garcia-Aymerich J, Serra I, Gomez FP, *et al.* Physical activity and clinical and functional status in COPD. *Chest* 2009; 136: 62–70.
47. Handschin C, Spiegelman BM. The role of exercise and PGC1alpha in inflammation and chronic disease. *Nature* 2008; 454: 463–469.
48. Pedersen BK. The diseasome of physical inactivity – and the role of myokines in muscle – fat cross talk. *J Physiol* 2009; 587: 5559–5568.
49. Langer D, Iranzo MA, Burtin C, *et al.* Determinants of physical activity in daily life in candidates for lung transplantation. *Respir Med* 2012; 106: 747–754.
50. Huppmann P, Sczepanski B, Boensch M, *et al.* Effects of in-patient pulmonary rehabilitation in patients with interstitial lung disease. *Eur Respir J* 2013; 42: 444–453.
51. Verlaet A, Moreira A, Sa-Sousa A, *et al.* Physical activity in adults with controlled and uncontrolled asthma as compared to healthy adults: a cross-sectional study. *Clin Transl Allergy* 2013; 3: 1.
52. Rasmussen F, Lambrechtsen J, Siersted HC, *et al.* Low physical fitness in childhood is associated with the development of asthma in young adulthood: the Odense schoolchild study. *Eur Respir J* 2000; 16: 866–870.
53. Vahlkvist S, Inman MD, Pedersen S. Effect of asthma treatment on fitness, daily activity and body composition in children with asthma. *Allergy* 2010; 65: 1464–1471.
54. Coronado M, Janssens JP, de Muralt B, *et al.* Walking activity measured by accelerometry during respiratory rehabilitation. *J Cardiopulm Rehabil* 2003; 23: 357–364.
55. Steele BG, Belza B, Hunziker J, *et al.* Monitoring daily activity during pulmonary rehabilitation using a triaxial accelerometer. *J Cardiopulm Rehabil* 2003; 23: 139–142.
56. Steele BG, Belza B, Cain KC, *et al.* A randomized clinical trial of an activity and exercise adherence intervention in chronic pulmonary disease. *Arch Phys Med Rehabil* 2008; 89: 404–412.
57. Dallas MI, McCusker C, Haggerty MC, *et al.* Using pedometers to monitor walking activity in outcome assessment for pulmonary rehabilitation. *Chron Respir Dis* 2009; 6: 217–224.
58. Mador MJ, Patel AN, Nadler J. Effects of pulmonary rehabilitation on activity levels in patients with chronic obstructive pulmonary disease. *J Cardiopulm Rehabil Prev* 2011; 31: 52–59.
59. Egan C, Deering BM, Blake C, *et al.* Short term and long term effects of pulmonary rehabilitation on physical activity in COPD. *Respir Med* 2012; 106: 1671–1679.
60. Sewell L, Singh SJ, Williams JE, *et al.* Can individualized rehabilitation improve functional independence in elderly patients with COPD? *Chest* 2005; 128: 1194–1200.
61. Mercken EM, Hageman GJ, Schols AM, *et al.* Rehabilitation decreases exercise-induced oxidative stress in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2005; 172: 994–1001.
62. Pitta F, Troosters T, Probst VS, *et al.* Are patients with COPD more active after pulmonary rehabilitation? *Chest* 2008; 134: 273–280.
63. Sallis JF, Prochaska JJ, Taylor WC. A review of correlates of physical activity of children and adolescents. *Med Sci Sports Exerc* 2000; 32: 963–975.
64. Springer AE, Kelder SH, Hoelscher DM. Social support, physical activity and sedentary behavior among 6th-grade girls: a cross-sectional study. *Int J Behav Nutr Phys Act* 2006; 3: 8.
65. O'Donnell DE, Casaburi R, Vincken W, *et al.* Effect of indacaterol on exercise endurance and lung hyperinflation in COPD. *Respir Med* 2011; 105: 1030–1036.
66. Hataji O, Naito M, Ito K, *et al.* Indacaterol improves daily physical activity in patients with chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis* 2013; 8: 1–5.
67. Cazzola M, MacNee W, Martinez FJ, *et al.* Outcomes for COPD pharmacological trials: from lung function to biomarkers. *Eur Respir J* 2008; 31: 416–469.

Chapter 11

Lung function



Denis E. O'Donnell, Pierantonio Laveneziana# and J. Alberto Neder**

SUMMARY: Pulmonary function tests are paramount for diagnosis, follow-up and prognosis in patients with suspected or established lung disease. Accordingly, they remain widely used in assessing the efficacy of medical and surgical treatments in respiratory medicine. Chronic obstructive pulmonary disease (COPD) epitomises a complex, multi-component disease in which pulmonary function tests have played a major role in establishing the current standards of treatment. However, it has become clear that assessment of the underlying physiological impairment in COPD requires an abundance of information, which exceeds that provided by simple spirometry. Based on the large accumulated experience from pulmonary function tests in clinical trials involving patients with COPD, this chapter provides an overview of the physiological rationale, strengths and limitations of a range of established and emerging physiological markers for potential use in future investigations of pharmacological therapy in mild-to-severe chronic respiratory disease.

*Dept of Medicine, Division of Respiratory and Critical Care Medicine, Queen's University, Kingston, ON, Canada. #Neurophysiologie Respiratoire Expérimentale et Clinique, ER 10 UPMC, Université Paris 6, Faculté de Médecine Pierre et Marie Curie (site Pitié-Salpêtrière), Paris, France.

Correspondence: D.E. O'Donnell, Dept of Medicine, Queen's University, 102 Stuart Street, Kingston, ON, K7L 2V6, Canada. Email: odonnell@queensu.ca

Eur Respir Monogr 2013; 62: 127–145.
Copyright ERS 2013.
DOI: 10.1183/1025448x.10045412
Print ISBN: 978-1-84984-044-6
Online ISBN: 978-1-84984-045-3
Print ISSN: 1025-448x
Online ISSN: 2075-6674

Pulmonary function tests remain fundamental tools for diagnosis, follow-up and estimation of prognosis in patients with lung disease. The physiological and technical basis of most of the currently available pulmonary function tests was established in the early 1950s. The available tests are sensitive to intervention and have acceptable degrees of reproducibility and external validity, making them useful instruments for the evaluation of therapeutic efficacy in clinical trials.

Chronic obstructive pulmonary disease (COPD) is a highly prevalent, chronic lung disease in which pulmonary function tests have played a decisive role in testing a number of alternative pharmacological treatments. COPD is characterised by heterogeneous inflammatory injury to the intrathoracic airways, lung parenchyma and pulmonary vasculature in highly variable combinations. Spirometric measurements, such as forced expiratory volume in 1 s (FEV_1), are required for diagnosis of COPD and are currently the principal physiological outcome used in clinical trials of bronchodilators, anti-inflammatory agents or a combination thereof [1]. Simple spirometry has stood the test of time but has recognised limitations which preclude its use as an exclusive outcome measure in clinical trials. For a given FEV_1 there is wide variability in airways resistance, pulmonary gas trapping, resting lung hyperinflation and the integrity of the alveolar–capillary gas exchanging interface [2]. These diverse physiological abnormalities ultimately underpin the clinical manifestations of COPD and are clearly inadequately captured by spirometry alone. For this reason, there is great interest in exploring additional or alternative physiological outcome measures, which may better characterise the response to pharmacotherapy in clinical trials involving these patients.

The objective of this chapter is to provide a concise overview of the rationale and performance characteristics of a number of emerging and established physiological measurements in clinical trials. We begin by briefly reviewing the strengths and limitations of traditional spirometry. We will then consider the clinical utility of various lung volume measurements used as primary or secondary outcomes in physiological studies and clinical trials in mild-to-severe COPD. Finally, we will review the physiological basis and potential role of measurements of peripheral airway function for future clinical trials in patients with early or spirometrically defined mild COPD.

Spirometry

Expiratory flow limitation is a defining physiological characteristic of COPD and represents the final expression of diverse derangements of respiratory mechanics. Since expiratory flow limitation is difficult to measure, spirometric measures of maximal expiratory flow rates are generally used in practice. Spirometry is effort dependent, influenced by inspiration volume history and subject to gas compression effects. Nevertheless, it is the most reproducible of all pulmonary function tests. Traditionally, an arbitrary increase in FEV_1 by 12% and 200 mL from baseline after short-term bronchodilator administration is considered to represent reversibility of airflow beyond the normal range of variability of the measurements [3]. An improved FEV_1 following bronchodilator administration indicates reduced resistance in the larger airways and alveolar units with faster mechanical time constants (*i.e.* the product of lung resistance and compliance) for lung emptying [4]. In more advanced COPD, an increased FEV_1 following a bronchodilator primarily indicates lung volume recruitment as a result of regional deflation of overinflated alveolar units (by reducing airflow resistance) [5–10]. Thus, in many cases, improvement in FEV_1 simply reflects recruitment of vital capacity (VC) as a result of reduced pulmonary gas trapping. This is reflected by a similar FEV_1/VC ratio pre and post-bronchodilator administration [5–10].

Historical concepts of reversibility in COPD have been challenged by a recent study by TASHKIN *et al.* [11], who have shown substantial acute improvements, beyond existing reversibility thresholds of FEV_1, following the combination of higher dose short-acting antimuscarinic and short-acting β_2-adrenoceptor agonist inhalation in patients with moderate-to-severe COPD enrolled in the UPLIFT trial [12]. Interestingly, in the cohort studied (n=5756), ~54% of patients showed a $\geqslant$12% and $\geqslant$200 mL increase in FEV_1, whereas ~66% showed a $\geqslant$15% improvement from baseline [11]. Thus, evidence of airway reversibility in "the asthmatic range" should not be considered an absolute exclusion criterion in clinical trials of pharmacotherapy in smokers with a history compatible with COPD.

Since forced vital capacity (FVC) measurement is time dependent and some patients with advanced COPD may not be able to sustain exhalation beyond 6 s, reliable values are more difficult to obtain. In this context, slow vital capacity (SVC) has been used as an alternative approach to circumvent the problem of FVC variability. FEV_6 has been shown to be more reproducible than FVC in a few clinical trials [13–16]. The trough (pre-bronchodilator) FEV_1 has been used as the primary outcome measure in recent placebo-controlled studies designed to test once or twice daily long-acting bronchodilators [17]. Improvements in FEV_1 >100 mL are regarded by regulatory authorities as being clinically acceptable, but this cut-off remains arbitrary [18].

Combinations of inhaled medications

Most, but not all, studies on the effects of combined short-acting [19, 20] and long-acting bronchodilators [21–24] confirm additive effects on FEV_1 compared with single drug components and placebo in moderate-to-severe COPD. Consistent amplification in bronchodilatation (by ~150 mL) has been found when inhaled corticosteroids (ICS; either fluticasone or budesonide) are added to long-acting β_2-adrenoceptor agonists (LABA; either salmeterol or formoterol) by mechanisms that are not fully understood [25–30]. Recent studies have examined the combined effects of triple therapy (*i.e.* a long-acting antimuscarinic, more commonly tiotropium, plus an

ICS/LABA combination) on airway function [21, 31–33]. Triple therapy was associated with a mean increase in FEV1 by ~100 mL and FVC by ~200 mL compared with each treatment alone in patients with more severe COPD [21, 31–33].

Measuring effect on disease progression

Traditionally, the rate of decline of post-bronchodilator FEV1 over time has been used to evaluate the impact of therapeutic interventions (*i.e.* smoking cessation, inhaled bronchodilators and ICS) on disease progression in COPD populations [12, 26, 34–38]. In most of these trials, FEV1 alone has failed to show a positive effect of the intervention under scrutiny. However, the lack of an independent criterion to define disease stability makes it difficult to ascertain whether this reflects the inadequacy of FEV1 or lack of efficacy of the intervention. In fact, ICS were not "consistently" associated with a slower rate of decline in FEV1 than placebo in patients with mild-to-severe COPD (ISOLDE: n=751, 3 years follow-up [36]; EUROSCOPE: n=1277, 3 years follow-up [37]; Copenhagen City Lung Study: n=290, 3 years follow-up [35]). Nevertheless, ~10–15% of COPD patients showed a significant post-bronchodilator FEV1 increase of at least 200 mL and 12% over baseline, which was maintained over time, when ICS was added to bronchodilator therapy in the EUROSCOPE study [37]. The addition of short-acting antimuscarinic inhalers to smoking cessation intervention in the Lung Health Study did not have additional effects on the rate of FEV1 decline [34]. Similarly, the UPLIFT study failed to show a reduction in the rate of FEV1 decline over 4 years in patients with moderate-to-severe COPD randomised to tiotropium or placebo [12]. Nevertheless, retrospective secondary analyses of the TORCH study have shown small but consistent reductions (compared with placebo) in the rate of FEV1 decline following treatment with ICS/LABA *versus* ICS alone and placebo [39].

Strengths

Low FEV measurements (<50% predicted) have been shown to correlate with poor health-related quality of life and mortality [40]. FEV measurements have been extensively standardised and are reproducible in the hands of well-trained technicians [41]. Reference population equations have been carefully derived and are widely available [42]. Spirometers are relatively inexpensive and the tests can usually be completed in a short time-frame. Therefore, spirometry is a safe, simple and effective tool for diagnosis and evaluation of treatment effects in clinical trials in COPD.

Limitations

In isolation, FEV1 is a poor index of severity of the complex multi-component disease that is COPD [40]. It is well established that smokers with a relatively preserved FEV1 can have extensive peripheral airways dysfunction, pulmonary gas exchange derangements, dyspnoea and activity restriction [2, 6, 43, 44]. FEV1 is less sensitive to pharmacological treatment in patients with more advanced COPD, in whom volume deflation effects are more pronounced, than in moderate COPD [10, 45]. Change in FEV1 after pharmacotherapy is generally not useful in predicting clinically meaningful improvement in dyspnoea, exercise intolerance, exacerbations and health-related quality of life. Although spirometry is favoured by regulatory authorities as a relevant outcome measure for use in trials, the minimal important difference (MID) for FEV1 and FVC measurements has not been established. It is now increasingly accepted that longitudinal decline in FEV1 may not provide a sufficiently comprehensive assessment of the effect of the pharmacological interventions on progressive respiratory impairment and its negative clinical consequences.

Lung volumes

Changes to the elastic properties of the lungs in emphysema alter the balance of forces between the lung (inward recoil) and chest wall (outward recoil) so that the relaxation volume of the respiratory system at the end of quiet expiration is re-set to a higher volume than is predicted in

health (*i.e.* resting "static" lung hyperinflation) [4, 46]. In the presence of expiratory flow limitation, expiratory time during spontaneous resting breathing is often insufficient to allow complete lung emptying and end-expiratory lung volume (EELV) is higher than the natural relaxation volume of the respiratory system [4, 46]. EELV is used interchangeably with functional residual capacity (FRC) in this chapter. Thus, in such patients EELV is, in part, dynamically determined. EELV can increase temporarily to a variable extent above the already increased resting value in situations where expiratory flow limitation is suddenly worsened (*i.e.* bronchospasm or exacerbation) or when ventilatory demand is abruptly increased, such as during exercise. This is termed dynamic lung hyperinflation [46]. The presence of resting and dynamic lung over-inflation reduces inspiratory capacity (IC) and, therefore, limits the ability of tidal volume (V_T) to expand during exercise, in turn contributing to early ventilatory limitation and intolerable dyspnoea (fig. 1) [46]. The dynamic component of lung hyperinflation, which is determined, in part, by autonomic control of airway smooth muscle tone, can be successfully manipulated by pharmacotherapy.

Decreasing lung hyperinflation

Decreased FRC following bronchodilation is explained by a reduction in the partially reversible, dynamically determined component of lung hyperinflation that is present in patients with resting expiratory flow limitation. Lung deflation is achieved by bronchodilation-induced airway smooth muscle relaxation, which accelerates mechanical time constants for lung emptying (by reducing airflow resistance) in heterogeneously distributed alveolar units [48]. Consequently, the lower operating lung volume and increased IC allow patients to achieve the required minute ventilation (V'_E) during rest and exercise at a lower oxygen cost of breathing [48]. Bronchodilators do not abolish expiratory flow limitation in patients with more advanced COPD but allow patients to generate the same or greater respired flow rates at lower operating lung volumes (fig. 2) [7–10, 49]. By deflating the lungs, they might improve ventilatory muscle performance leading to greater V_T expansion as V'_E increases during exercise [5, 7–9, 48, 49]. Thus, neuromechanical coupling is enhanced and dyspnoea is lessened [48]. Consequently, patients can exercise for longer before the mechanical limitation to V'_E is reached [5, 7–9, 48, 49].

Short-term improvements in residual volume (RV) and FRC following treatment with single inhaled long-acting bronchodilator agents average 450 mL (~10% from baseline) and 400 mL (~8% from baseline), respectively [8, 9, 49, 51]. CHRISTYN *et al.* [52] were the first to show a dose–response relationship between bronchodilator (oral theophylline) administration and reduction in FRC.

Inspiratory capacity

There has been considerable interest in measuring the acute change (increase) in spirometric IC as a measure of lung deflation following bronchodilator treatment in COPD. The resting IC indicates the position of V_T relative to total lung capacity (TLC) on the sigmoidal pressure–volume relationship of the relaxed respiratory system. Therefore, an increase in resting IC following bronchodilation indicates reduced elastic/inspiratory threshold loading of the inspiratory muscles, an important determinant of dyspnoea [30, 53, 54]. Post-bronchodilator change in resting IC of ~300 mL or 10–15% pred (or 15–17% of baseline value) has been consistently associated with increased exercise endurance time in the order of 20–30% in patients with moderate-to-severe COPD (fig. 3) [5–9, 48, 49, 51, 55, 56].

Average changes in resting IC following single long-acting bronchodilator therapy range from 200 to 400 mL [8, 9, 49, 51, 55]. However, it should be noted that this improvement in IC underestimates the true reduction in FRC by the amount by which the TLC falls (*i.e.* ~120 mL on average) [8, 9, 49, 51, 55]. Despite these concerns, intra-individual differences in plethysmographically determined FRC and EELV (TLC minus IC) and their response to bronchodilators have not been specifically addressed in a large number of patients with COPD.

The largest improvements in IC following bronchodilators are seen in patients with the greatest resting lung hyperinflation (*e.g.* baseline IC <80% pred) [10, 45, 57]. VAN NOORD *et al.* [24] were the first to study the combined effects of two long-acting bronchodilators (tiotropium and formoterol) on lung deflation over a 24-h period in patients with moderate-to-severe COPD. They confirmed additive effects on lung deflation with average increases in IC of 220 mL and daytime peak increase of 550 mL. Decreases in lung volume of this magnitude are associated with reduced intrinsic mechanical loading and increased functional strength of the respiratory muscles [48]. Several studies have shown a preserved or even decreased FEV_1/FVC ratio in response to all classes of bronchodilators, *i.e.* greater changes in FVC than FEV_1 [7–10, 45, 49, 51, 55, 56]. This pattern of lung volume recruitment is particularly prominent in patients with severe lung hyperinflation [10, 45, 49]. It is noteworthy that the primary effect of bronchodilator therapy is on the resting IC (reducing lung hyperinflation), which increases the available dynamic inspiratory reserve volume and delays the critical respiratory mechanical limitation to exercise. Accordingly, the rate of post-bronchodilator dynamic lung hyperinflation is generally not decreased during exercise (or may actually increase), reflecting the increased V'_E secondary to the release of V_T restriction due to less lung hyperinflation. Nevertheless, measurements of sub-maximal exercise tolerance and dyspnoea add important information to resting IC as the multifactorial nature of exercise intolerance in COPD makes the impact of resting lung deflation on patients' functional performance unpredictable [5–9, 48, 49, 51, 55, 56].

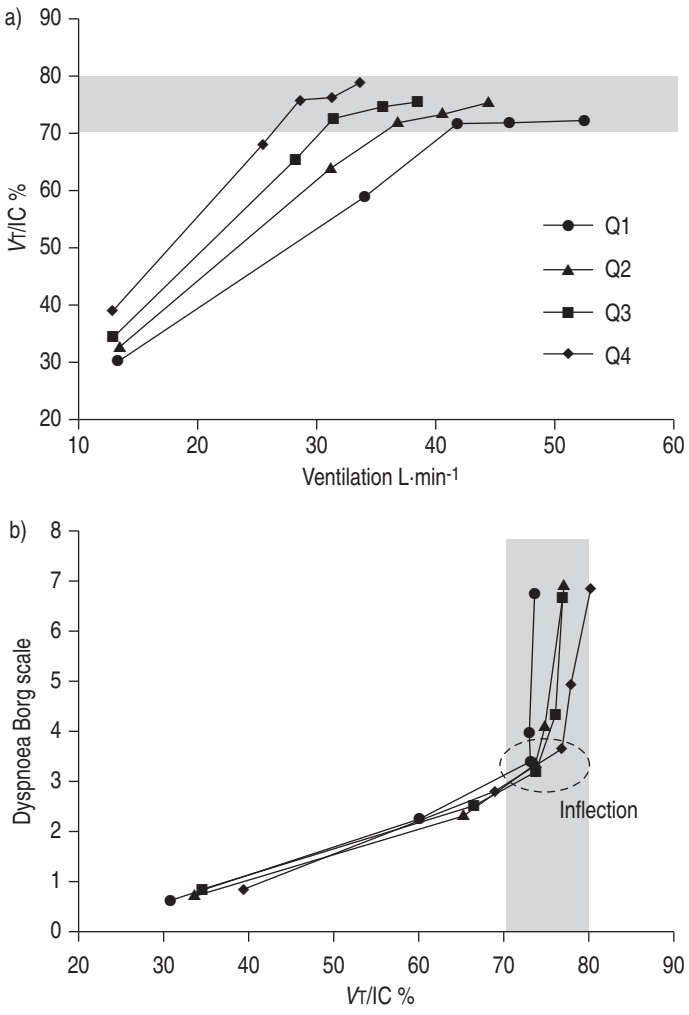


Figure 1. a) In hyperinflated patients with chronic obstructive pulmonary disease (COPD), tidal volume (V_T) expands to a maximal value of 70–80% of the actual inspiratory capacity (IC). b) When this inflection/plateau in V_T/IC occurs, dyspnoea rises sharply to intolerable levels, regardless of baseline differences in lung function. The quartiles of COPD severity (airflow limitation) were as follows. Q1: forced expiratory volume in 1 s (FEV_1) 62% pred; Q2: FEV_1 49% pred; Q3: FEV_1 39% pred; Q4: FEV_1 27% pred. Reproduced from [47] with permission from the publisher.

Studies that have investigated the effects of short- and long-acting bronchodilators on lung hyperinflation during exercise are summarised in table 1. The results demonstrate improved IC throughout rest and exercise, in association with improvement in dyspnoea intensity ratings during exercise, and increased exercise endurance (fig. 2) [5–9, 22, 48, 49, 51, 55, 56, 58–65, 67–75]. Some studies have compared the effects of long-acting bronchodilators. Tiotropium was reported to be superior to salmeterol in improving resting volumes and exercise endurance [74, 75]. Similarly, indacaterol was reported to be superior to salmeterol for its effect on resting IC [67].

Combining two bronchodilators generally extends lung deflation improvements seen with single agents, *e.g.* resting IC was increased with indacaterol plus tiotropium compared with tiotropium alone [76], or with formoterol (or its isomer arformoterol) plus tiotropium *versus* tiotropium alone [24, 77, 78]. The addition of tiotropium to an ICS/LABA combination has been reported to improve resting lung volumes compared with either individual component [31–33]. The bronchodilator combination of tiotropium and salmeterol was more effective than an ICS/LABA for improving resting lung volumes in hyperinflated patients, although exercise endurance time was not significantly increased [79].

ICS and ICS/LABA combinations

O'Donnell *et al.* [55] reported that fluticasone/salmeterol reduced resting and dynamic hyperinflation, and increased exercise endurance time compared with placebo but not *versus* salmeterol. A later study found that adding fluticasone to long-acting bronchodilator therapy decreased EELV during exercise without consistent improvements in dyspnoea intensity and exercise endurance [53]. Budesonide/formoterol was reported to increase exercise endurance and reduce dynamic hyperinflation compared with both placebo and formoterol [66]. The mechanism for an effect of ICS on hyperinflation is unclear but may involve a direct local action on pulmonary vasculature, *i.e.* vasoconstriction [53, 80, 81]. Beclomethasone/formoterol was reported to improve resting lung volumes more than fluticasone/salmeterol [82]. In addition, budesonide/formoterol showed a faster rate of onset than fluticasone/salmeterol [82] or formoterol alone [83].

Effect of other treatments on lung hyperinflation

Theophylline has been reported to improve lung volumes and exercise endurance [84, 85], possibly through improved respiratory muscle performance secondary to unmeasured lung volume reduction [86, 87]. The oral phosphodiesterase-4 inhibitor, roflumilast, was associated with progressive improvement of airway function but not resting lung hyperinflation compared with placebo after 12 weeks of treatment [88]. Roflumilast had small but consistent effects on air trapping during exercise (IC increased by 120 mL ($p=0.008$) at iso-time compared with placebo), but this was not sufficient to improve dyspnoea or exercise endurance time [88]. Improved hyperinflation and exercise endurance were reported following treatment with the mucoactive/antioxidant treatment acetylcysteine [89].

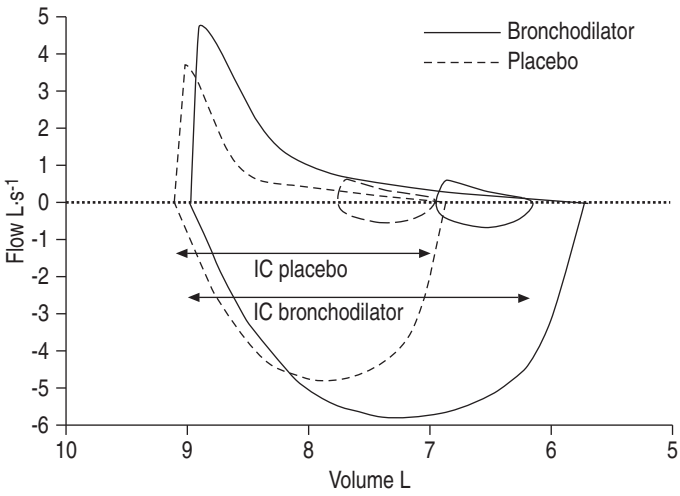


Figure 2. In a typical patient with severe chronic obstructive pulmonary disease (COPD), maximal expiratory flows increased from placebo (– – – –) after a bronchodilator (——) in association with a decrease in end-expiratory lung volume, as reflected by an increase in inspiratory capacity (IC). The bronchodilator did not abolish expiratory flow limitation in patients with more advanced COPD but allowed tidal volume (inner loops) to be positioned at a lower operating lung volume. Reproduced from [50].

Decreasing lung hyperinflation in mild COPD

Despite the largely preserved FEV_1, a number of recent studies have shown that symptomatic patients with mild COPD (as spirometrically defined) had evidence of extensive small airway dysfunction (see Tests of small airways function section), higher ventilatory requirements throughout exercise and dynamic pulmonary gas trapping as a result of expiratory flow limitation [47, 53, 54, 90, 91]. In small physiological studies, bronchodilator or ICS/LABA therapy (compared with placebo) was

associated with consistent reduction of pulmonary resistance, lung hyperinflation, and work and oxygen cost of breathing during exercise but with no consistent improvement in dyspnoea or exercise endurance in mild COPD [53, 54].

Strengths

Lung volume measurements are clinically relevant and provide complementary information to spirometry for the individual assessment of the COPD patients. Plethysmographic measurements of lung volume components are well standardised and published reference values are widely available [3, 41, 92]. The plethysmographic FRC measurement is highly reproducible with a coefficient of variation ranging from 3.5% to 6.7% [93]. The IC/TLC ratio has been shown to be an independent predictor of all-cause and respiratory mortality in severe COPD [94]. Data from one clinical trial indicated good repeatability and responsiveness of IC, TLC, FRC and RV measurements following bronchodilator treatment [8]. Reduced resting lung hyperinflation following bronchodilator treatment has been consistently associated with clinically important improvements in dyspnoea and exercise tolerance in a number of recent clinical trials (table 1).

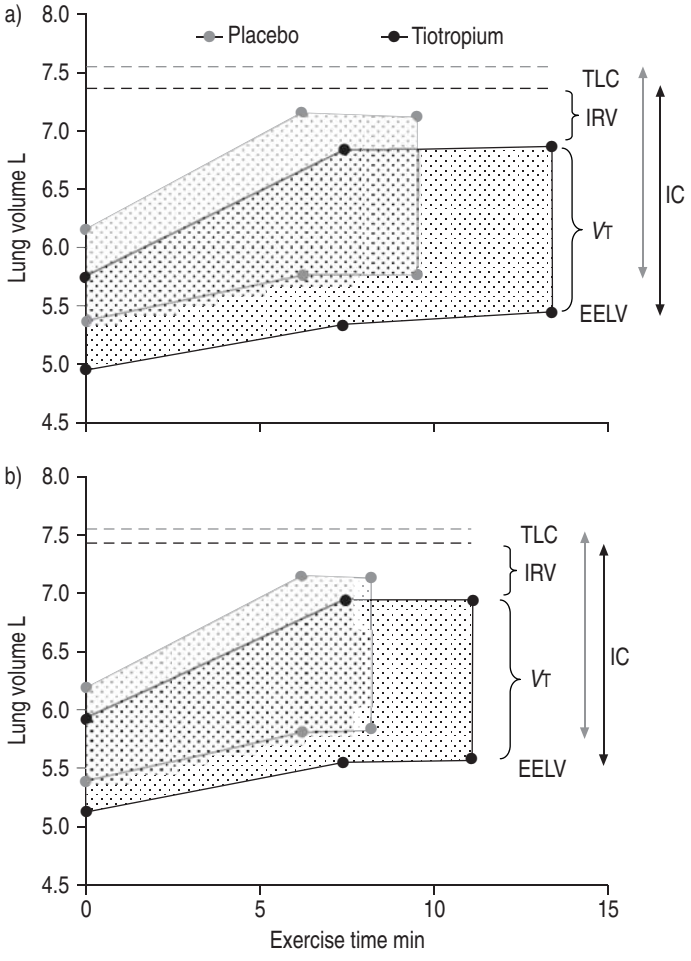


Figure 3. Operating lung volumes during constant work rate cycle ergometry to symptom limitation at 75% of maximal work capacity on day 42 of treatment: a) 2.25 h post-dose and b) 8 h post-dose. At 2.25 h post-dose, total lung capacity (TLC) corresponds to body plethysmography measurements at 80 min post-dose. At 8 h post-dose, TLC corresponds to values calculated by linear interpolation using values at trough (pre-dose) and 80 min post-dose. IRV: inspiratory reserve volume; *V*T: tidal volume; EELV: end-expiratory lung volume; IC: inspiratory capacity. Reproduced from [9] with permission from the publisher.

Limitations

The natural history of lung hyperinflation and its clinical implications for individual patients with COPD is unknown. No consensus exists regarding stratification of severity of lung hyperinflation. Plethysmographic lung volume measurements are more expensive, time consuming and technically demanding than spirometry. Reproducibility data on absolute lung measurements are lacking as are population reference values for IC. The latter are currently derived as the differences of predicted TLC and FRC reference values. Ideally, the use of reduced IC as a measure of lung hyperinflation in COPD requires baseline resting plethysmographic lung volumes to rule out concomitant restrictive lung disease (decreased chest wall compliance, lung fibrosis or respiratory muscle weakness). In overweight and obese COPD patients IC is actually preserved or increased compared to lean patients with a similar FEV_1 [80]. As with other spirometric indices,

Table 1. Effect of long-acting bronchodilators, inhaled corticosteroids, and inhaled corticosteroids + long-acting β_2-agonists on resting lung volumes and exercise variables

First author [ref.]	Treatment dosage	Duration	FEV_1 % pred	Resting measurements post-dose			Exercise measurements post-dose		
				FRC L or % change	RV L or % change	IC L or % change	IC L or % change	Borg dyspnoea score	Endurance time
O'Donnell [6]	IB 500 μg Placebo	Beginning of 3 weeks	40±2	NA	NA	3 weeks: 0.45* from baseline	3 weeks Peak: 0.41* from baseline	3 weeks Isotime: -0.6* from baseline	168 s* from baseline
		End of 3 weeks		NA	NA	3 weeks: 0.33* from baseline	3 weeks Peak: 0.30* from baseline	3 weeks Isotime: -0.3* from baseline	3 weeks 66 s* from baseline
Laveneziana [58]	IB+Salb 500/2500 μg	Acute administration	57	-0.40*	-0.50*	0.40*	Peak: 0.40	Isotime: -0.5	66 s
O'Donnell [48]	Tio 18 μg once daily Placebo	7–10 days	40±3	-0.29*	-0.34*	0.20*	Peak: 0.31*	Isotime: -0.7#	54 s
O'Donnell [8]	Tio 18 μg once daily Placebo	6 weeks	41–42	Trough: -0.30* Peak: -0.41*	Trough: -0.36* Peak: -0.51*	Trough: 0.10* Peak: 0.25*	Peak: 0.15*	Isotime: -0.9*	105 s*
Maltais [9]	Tio 18 μg once daily Placebo	6 weeks	43	Trough: -0.17* Peak: -0.40*	Trough: -0.29* Peak: -0.54*	Trough: 0.17* Peak: 0.22	Peak: 0.14* to 0.13* (2 and 8 h post-dose)	Isotime: -1.05* to -0.97* (2 and 8 h post-dose)	236 s* to 171 s* (2 and 8 h post-dose)
Maltais [59]	Acl 200 μg once daily Placebo	6 weeks	49–52	-0.40*	NA	0.3*	Peak: 0.02	Isotime: -0.7*	116 s*
Beeh [60]	Glycopyrronium 50 μg once daily Placebo	3 weeks	57	-0.46*	-0.50*	0.22–0.23*	Peak: 0.19*	Isotime: -1.16*	89 s*
Berton [61]	For 12 μg twice daily + Tio 18 μg once daily	2 weeks	47	NA	NA	For+Tio: 0.23	Peak: NA	For+Tio Isotime: -2	For+Tio: Improvers: 548 s* Non-improvers: 194 s*
	For 12 μg twice daily + placebo			NA	NA	For+placebo: 0.26	Peak: NA	For+placebo: Isotime: -2	For+placebo: Improvers: 236 s Non-improvers: 377 s
Neder [62]	For 12 μg twice daily Placebo	2 weeks	39	-0.29*	NA	0.28*	Peak: 0.14*	Isotime: -1.5	130 s
O'Donnell [49]	Slm 50 μg twice daily Placebo	2 weeks	42	-0.35*	-0.43*	0.33*	Peak: 0.22*	Isotime: -0.9	96 s*
Beeh [63]	Ind 300 μg once daily Placebo	2 weeks	52 (post-BD)	-0.34	NA	0.18*	Peak: 0.32*	Isotime: -1.5*	88 s*
O'Donnell [64]	Ind 300 μg once daily Placebo	3 weeks	61 (post-BD)	NA	NA	Pre-dose: 0.14* Post-dose: 0.17*	End-exercise: 0.28*	NA	111 s*
Guenette [65]	FP 500 μg twice daily Placebo (added to BD treatment)	2 weeks	54	-0.30*	-0.32*	0 (both 2.71)	Peak: -0.04	Isotime: -0.7	Peak: 188 s*
Worth [66]	Bud/For 320/9 μg twice daily	1 week	38	Bud/For: -4%*	Bud/For: -4%*	Bud/For: 14%*	NA	Isotime 1 h post-dose:	1 h post-dose:
				For: -1%	For: -1%	For: 12%*		Bud/For: -0.8*	Bud/For: 105 s*
				Bud/For *versus* For: -3%	Bud/For *versus* For: -3%	Bud/For *versus* For: 2%		For: approx. -0.3	For: 36 s
								Bud/For *versus* For: approx.-0.5	Bud/For *versus* For: 69 s*
	For 9 μg twice daily			Bud/For: -5%*	Bud/For: -5%*	Bud/For: 14%*	NA	Isotime 6 h post-dose:	6 h post-dose:
	Placebo			For: -3%	For: -3%	For: 9%*		Bud/For: -1.1*	Bud/For: 62 s*
				Bud/For *versus* For: -2%	Bud/For *versus* For: -2%	Bud/For *versus* For: 5%		For: -1.1*	For: 20 s
								Bud/For *versus* For: approx. 0	Bud/For *versus* For: 42 s*

Table 1. Continued.

First author [ref.]	Treatment dosage	Duration	FEV1 % pred	Resting measurements post-dose			IC L or % change	Exercise measurements post-dose	
				FRC L or % change	RV L or % change	IC L or % change		Borg dyspnoea score	Endurance time
O'Donnell [55]	FP/Slm 250/50 µg twice daily	8 weeks	40–43	FP/Slm: -0.35*	FP/Slm: -0.35*	FP/Slm: 0.33*	FP/Slm Peak: 0.17*	FP/Slm Isotime: -0.5	FP/Slm: 132 s*
	Slm 50 µg twice daily			Slm: -0.26*	Slm: -0.30*	Slm: 0.27	Slm Peak: 0.11	Slm Isotime: -0.4	Slm: 86 s
	Placebo								
Guenette [53]	FP/Slm Placebo	6 weeks	86±15	Trough: -0.15* Peak: -0.28*	Trough: -0.06 Peak: -0.17*	Trough: 0.15* Peak: 0.18*	Peak: 0.17*	Isotime: -0.4	72 s

FEV1: forced expiratory volume in 1 s; % pred: % predicted; FRC: functional residual capacity; RV: residual volume; IC: inspiratory capacity; IB: ipratropium bromide; NA: not available; Salb: salbutamol; Tio: tiotropium; Acl: aclidinium; For: formoterol; Slm: salmeterol; Ind: indacaterol; FP: fluticasone propionate; Bud: budesonide; BD: bronchodilator. *: $p<0.05$ *versus* placebo; #: $p=0.06$.

no MID has been established for direct or indirect measurements of lung hyperinflation. Compared with spirometry, experience with lung volume outcomes in bronchodilator trials is limited.

Tests of small airway function

It has long been recognised that FEV1 is insensitive to functional changes of the small airways, *i.e.* those <2 mm in diameter [95]. These peripheral airways have long been named the lung's "quiet zone" [96] as it has been estimated that obstruction of up to 75% of peripheral airways is required before changes can be detected by FEV1. The small airways are non-cartilaginous bronchioles which delineate the transition from purely conducting airways to gas exchanging tissue [97]. Deposition of suspended particles is particularly increased in the small airways owing to the sudden slowing of flow as their total cross-sectional area increases rapidly. Thus, they receive the greatest burden of inhaled particles per unit of surface area [98]. The prevailing view is that the recurrent inflammatory response to particles deposition and/or to smoking irritation at the level of the small airways start early in the natural history of COPD [44, 99]. Wall thickening, cellular infiltration and occlusion by exudates then contribute to increase small airway resistance. Destruction and loss of terminal bronchiole ensues and coalescence of several primary lesions forms larger centrilobular emphysematous spaces. As emphysema progresses, loss of alveolar attachments potentiates small airway closure at low lung volumes, particularly in the dependent lung regions [44]. It has been proposed that mechanical injury induced by small airways closure during tidal breathing may also worsen airway inflammation which, in turn, may accentuate peripheral airway injury with the results of promoting and accelerating the abnormalities of lung function in a vicious circle [100].

As expected by their anatomical and functional properties, small airways (dys)function is notoriously difficult to assess [101]. As previously mentioned, comparative studies of spirometry and other tests of small airway function consistently found that FEV1 was the least sensitive functional parameter for measuring the physiological effects of bronchodilators in COPD. However, morphological analysis of maximal flow–volume curves might be instructive as regional heterogeneity of flow–time constants, progressive increases of airways resistance with lung deflation and premature airway closure may create an upper concavity in the expiratory limb ("scooping"). Unfortunately, there is

currently no systematic method to quantify such abnormalities. Mid-expiratory flows (*e.g.* forced expiratory flow at 25–75% of FVC ($FEF_{25–75\%}$)) can also increase the sensitivity of spirometry to detect early obstruction; however, there is a marked measurement variability (lower limit of normality ~60% pred) [3]. $FEF_{25–75\%}$ also varies in proportion to FVC thereby widely shifting the lung volume range at which the $FEF_{25–75\%}$ is computed [102]. $FEF_{25–75\%}$ has been found to correlate poorly with measures of air trapping [103, 104] and has added little to early detection of COPD in large epidemiological studies [104, 105]. Nevertheless, COPD patients with a history of frequent exacerbations showed decreased $FEF_{25–75\%}$ values which selectively improved with tiotropium [106].

To date, although the so-called "tests of small airways function" [107] have not yet been incorporated in large clinical trials, there is emerging evidence that they might uncover physiological benefits of bronchodilators which are not detected by spirometry. This is particularly relevant for patients with mild COPD who are more likely to present with a "positive" response to these medications when a more sensitive test is used [108]. The small airways are usually evaluated indirectly, taking advantage of the functional consequences of: 1) premature airway closure and air trapping; 2) regional heterogeneity of mechanical time constants due to variations in either the compliance or resistance of alveolar units; 3) delayed gas emptying in alveolar units that are slowly ventilated by collateral channels; and 4) exaggerated volume dependence of airflow limitation (table 2) [101, 109–111].

Tests of airways closure

Physiological rationale

Exaggerated dependency on lung volume and transmural pressure gradients leading to early small airway closure and air trapping can be assessed with some tests that are familiar to most clinical laboratories. For instance, FVC can be substantially reduced (>200 mL) compared to SVC when the small airways close at low lung volumes during the forced manoeuvre [3]. In fact, the

Table 2. Selected tests of small airways function and their interpretation in the presence of small airways disease

Pathophysiological feature	Test	Main parameters	Expected result
Early airways closure	Spirometry	FVC	↓
		FVC/SVC	↓
		$FEF_{25–75\%}$	↓
	Body plethysmography	RV	↑
		RV/TLC	↑
	Single-breath nitrogen washout	Closing volume	↑
		Closing capacity	↑
Increased airways resistance	Body plethysmography	sG_{aw}	↓
	Heliox *versus* room air	Flow at 50% VC	↔
	Bronchial hyperresponsiveness	PC_{20}	↓
	Oscillometry	X_{rs} 5–20	↑
		X_{rs} 5	↑
		Resonant frequency	↑
Ventilation inhomogeneity	Single-breath nitrogen washout	Phase III slope	↑
	Multiple-breath nitrogen washout	Lung clearance index	↓
		S_{cond}	↓
		S_{acin}	↓ or ↔

FVC: forced vital capacity; SVC: slow vital capacity; $FEF_{25–75\%}$: forced expiratory flow at 25–75% of FVC; RV: residual volume; TLC: total lung capacity; sG_{aw}: specific airway conductance; VC: vital capacity; PC_{20}: provocative concentration causing a 20% fall in forced expiratory volume in 1 s; X_{rs}: respiratory system reactance; S_{cond}: index of conductive ventilation inhomogeneity; S_{acin}: index of acinar ventilation inhomogeneity; ↑: increased; ↓: decreased; ↔: unaltered.

difference between inspiratory SVC and FVC and FVC/SVC ratio have been used as surrogate markers of small airways collapsibility as they correlated well with more complex tests of small airways function [112]. FVC has been found to correlate inversely with RV/TLC in patients with severe, poorly reversible asthma who are functionally indistinguishable from COPD [113]. A decrease in FVC at the provocative concentration causing a 20% fall in FEV_1 has also been assumed as indicative of air trapping [114]. Other physiological measurements influenced by small airways closure at low lung volumes include RV and RV/TLC ratio by body plethysmography and closing volume/capacity by single-breath nitrogen washout [109–111]. In the later test, expired nitrogen concentration is measured after a single inhalation of 100% oxygen. The sudden increase in expired nitrogen concentration near the end of expiration marks the volume at which the small airway starts to close, a finding that is positively related to structural small airways abnormalities [109–111].

Effects of pharmacological interventions

As discussed previously (see Lung volumes section), bronchodilator-induced improvements in lung hyperinflation (*e.g.* larger IC and lower EELV) can positively impact on exercise tolerance and dyspnoea in COPD [5–9, 48, 49, 51, 55, 56]. Other authors used the expected inverse relationship between FVC and pulmonary gas trapping to show the positive effects of *N*-acetylcysteine [89] and an extra-fine combination of beclomethasone and formoterol [115] on small airway function in patients with COPD. Importantly, these positive findings were largely independent of drug-induced changes in FEV_1. TZANI *et al.* [81] demonstrated less pulmonary gas trapping (reduced RV and TLC) with the beclomethasone/formoterol extrafine combination than with standard-size fluticasone/salmeterol combination in hyperinflated patients with COPD. Reductions in static lung volumes, despite the lack of change in FEV_1, were also demonstrated in the BRONCUS trial with *N*-acetylcysteine [116].

Tests of airways resistance

Physiological rationale

In contrast to health, the predominant site of airway narrowing during tidal breathing seems to be located at the small airways in patients with COPD. Loss of static lung recoil pressure enhances expiratory dynamic compression, especially in the non-cartilaginous airways [105, 111]. Body plethysmography takes advantage of these phenomena to measure airways resistance (R_{aw}; pressure/flow) and its reciprocal (airway conductance, G_{aw}) during tidal breathing. G_{aw} corrected for the specific volume at which it is measured (sG_{aw}) has been shown to correlate moderately well with histopathological abnormalities in the small airways of COPD [117]. Another method to investigate R_{aw}, which is more specific to small airways, compares maximum expiratory flows within room air and a low-density gas mixture (a helium-oxygen mixture, usually 79% helium and 21% oxygen, also called heliox). As airflow patterns are typically laminar in the small airways and turbulent in larger airways, a lack of effect of heliox on expiratory flow rate (*e.g.* at mid expiration) suggests that the small airway is the major site of resistance [109–111]. The small airway also seems to be particularly responsive to methacholine provocation, a phenomenon demonstrated directly by wedged bronchoscopy [114] and indirectly by air trapping on computed tomography [118], worsening of ventilation/distribution inequalities [119], and distal heterogeneity by oscillometry [120]. Of note, airway hyperresponsiveness was inversely related to airway wall thickness in resected lungs after accounting for lung elastic recoil and FEV_1 % pred in patients with mild-to-moderate COPD [121].

Oscillometry has gained increasing popularity as a method to assess small airway resistance [105, 111]. Small pressure oscillations (usually 5–20 Hz) are applied at the mouth during tidal breathing to obtain a measure of respiratory impedance. Multiple frequencies can be applied in the form of a pseudo-random, continuous mixture of several sinusoidal waveforms (forced oscillation technique (FOT)) or by rapidly recurring rectangular waveform impulses (impulse oscillometry) [122]. Impedance of the respiratory system is the ratio of pressure to flow and is

composed of resistance (R_{rs}) and reactance (X_{rs}) of the respiratory system. The frequency at which X_{rs} equals zero is the resonant frequency. R_{rs} is a reflection of the frictional pressure losses as air travels in and out of airways and it is directly related to airway calibre. X_{rs} reflects not only the elastic properties of the airways but is also affected by inhomogeneity in ventilation thereby reflecting frequency dependence of compliance, *i.e.* the expected reduction in lung compliance when there is substantial inequality in the speed at which the airways empty [123].

It is widely believed that low-frequency R_{rs} ($\leqslant$10 Hz) reflects the small conducting airways (for frequencies <10–15 Hz the oscillation pressures penetrate into the peripheral lung), and high-frequency R_{rs} preferentially represents the proximal airways (for frequencies >15 Hz the oscillation pressures are absorbed before reaching the small conducting airways), while X_{rs} reflects the mechanical properties of the small airways [122]. COPD patients usually present with frequency dependence of resistance as the resistance at low frequencies (*e.g.* at 5 Hz) is raised to a greater extent than that at high frequencies (*e.g.* at 20 Hz) [124–129]. In COPD, oscillometry variables, particularly X_{rs}, have been successfully used to distinguish healthy adults from those with COPD [130], detect early changes in lung mechanics in smokers [131], identify patients with combined pulmonary fibrosis and emphysema [132], differentiate COPD from asthma [133], establish the presence of resting expiratory flow limitation [134], predict the extension of emphysema [126] and track functional recovery after exacerbations [135]. Moreover, impulse oscillometry measurements correlated better with the severity of chronic activity-related dyspnoea (assessed using the Medical Research Council scale) than FEV_1 [136] and predicted bronchodilator-induced improvements in exertional dyspnoea induced by a 6-min walking test [137]. Of note, X_{rs} and R_{rs} have been found to be highly repeatable in COPD subjects [138], even when unsupervised measurements were obtained at home [139].

Effects of pharmacological interventions

A number of studies showed that R_{rs} and X_{rs} at low frequencies (*e.g.* at 5 Hz) and resonant frequency were all more sensitive than FEV_1 to detect the effects of bronchodilators in COPD [127–129, 140, 141]. For instance, DELLACÀ *et al.* [140] identified bronchodilator responders to inhaled salbutamol (improvement in inspiratory resistance at 5 Hz) with a discriminative power of 80%. DI MANGO *et al.* [128] reported that acute inhalation of salbutamol was associated with reductions in the parameters associated with total and airway resistance, as well as improvement in homogeneity and dynamic compliance of the respiratory system. More recently, GUENETTE *et al.* [53] found that administration of inhaled fluticasone/salmeterol significantly reduced impulse oscillometry measurements at low frequencies, such as resistance at 5 Hz, reactance at 5 Hz and resonant frequency, compared with placebo, while measurements at higher frequencies (*i.e.* 15 and 20 Hz) did not change.

Important studies by BORRILL and co-workers [127, 129] demonstrated that FOT detected improvement in lung function following lower doses of bronchodilators, and was able to distinguish between different classes of these medications in the absence of spirometric changes in patients with COPD. In the first study [127], the authors reported that whereas reactance at 5 Hz, sG_{aw} and resistance at 5 Hz detected significant effects of salbutamol at doses as low as 20 μg (ranging from 11.5% to 21.5% change), FEV_1 was the least sensitive parameter, detecting significant bronchodilation only at 100 μg (10.2% change). In the second study [129], they reported that despite similar positive effects on FEV_1, resistance at 35 Hz showed superiority of salmeterol over tiotropium within the first 12 h post-dose, whereas sG_{aw}, resistance at 5 Hz, resistance at 15 Hz and resonant frequency demonstrated superiority of tiotropium over salmeterol from 12 h to 26 h. More recently, TSE *et al.* [141] reported significant improvements in FOT parameters after 1 year of high-dose *N*-acetylcysteine in a double-blind placebo-controlled trial involving chronic bronchitis patients. Body plethysmographic measurements are more widely available than impulse oscillometry and FOT and they have recently been considered by authoritative reviews as promising measurements to be included in large trials

[101, 105, 111]. However, R_{aw} and specific R_{aw} cannot be regarded as specific tests of small airway dysfunction.

Tests of ventilation distribution homogeneity

Physiological rationale

In the single-breath nitrogen washout test, the expired nitrogen concentration slope against expired volume (phase III) provides an index of the overall inhomogeneity of ventilation [105, 111]. Multiple-breath nitrogen washout tests have long been used to estimate FRC [142]. The rate at which nitrogen is washed out by oxygen can also provide an index of gas mixing efficiency. In the simplest form, this can be estimated by recording the number of breaths necessary to reach a pre-determined expired nitrogen concentration. More elaborate strategies include the calculation of a lung clearance index (LCI), which represents the necessary ventilation (corrected by lung size using FRC as reference) to reduce expired nitrogen concentration to a given value (usually 1/40th of the starting concentration) [142]. For instance, LCI has been found to be more sensitive than other measures of lung function in detecting small airway abnormalities by high-resolution computed tomography [143].

More recently, a new multiple-breath nitrogen washout approach has been proposed to assess topographical variability of gas mixing within an airways zone in which airflow is convection-dependent (*i.e.* dependent on pressure gradients) and another beyond the point at which gas flow becomes diffusion dependent (*i.e.* dependent on concentration gradients) [144, 145]. In this method, the nitrogen phase III slope is computed and normalised by the mean expired nitrogen concentration leading to a normalised slope for each breath. Assuming that the rate of rise of the normalised slope curve is due to the convective flow asynchrony between lung units larger than acini, and the offset of the curve is mainly determined by diffusion-heterogeneity in the acinar airways, the method provides an index of conductive ventilation inhomogeneity (S_{cond}) and an index of acinar ventilation inhomogeneity (S_{acin}) [144, 145].

Effects of pharmacological interventions

The single- and multiple-nitrogen breath washout tests have been used to assess the effects of bronchodilators in asthma and cystic fibrosis [146, 147], particularly in the paediatric group. However, in patients with COPD their use in pharmacological trials are still very limited. For instance, phase III slope by single-breath nitrogen washout correlated significantly with exhaled nitric oxide and showed superior responses to fluticasone/salmeterol compared with FEV_1 in patients with moderate COPD [148]. Evaluating the effects of smoking cessation in mild-to-moderate COPD, other studies found that despite the lack of change in FEV_1 and $FEF_{75\%}$, there were substantial increases in S_{cond} only 1 week after quitting [101]. In addition, this improvement was sustained after 1 year and a 42% reduction of conductive airway abnormality was observed [101]. These data suggest that the multiple-breath nitrogen washout tests may expose functional improvements following pharmacological treatment in COPD well before and at a greater extent than the traditional tests.

Strengths

The small airways are affected early in the natural history of COPD [101, 105, 107–111] and evidence of small airways dysfunction can be identified well before the development of overt spirometric abnormalities [13, 14, 28, 69, 106, 116]. Increased test sensitivity can be instrumental in showing subtle differences between the action of closely related pharmacological treatments, *e.g.* between different long-acting bronchodilators. There is growing interest in the development of formulations targeted at the peripheral airways thereby anticipating that these tests will be more frequently used in the future. Importantly, most tests of small airway function did not require maximal respiratory manoeuvres making them more externally valid and less effort-dependent than the traditional tests. In fact, they usually relate better to patient-reported outcomes than spirometry and static lung

volumes [13, 14, 28, 69, 106, 116]. Nevertheless, it should be acknowledged that these positive results in favour of the tests of small airway function might be influenced by some reporting bias of small physiological studies performed in highly specialised laboratories.

Limitations

The tests of small airway function have not been incorporated into clinical practice more than 30 years after their conception which suggests important limitations to their widespread application. For instance, methodological approaches to measure structural abnormalities in the small airways are still lacking. This indicates that our ability to unequivocally link the functional effects of a given pharmacological approach with small airway treatment remains rather limited. The tests of small airway dysfunction are notoriously less reproducible than spirometric measurements [13, 14, 28, 69, 106, 116] and results from different tests may not agree with each other [13, 14, 28, 69, 106, 116]. Age- and sex-based reference values are still not available and, in most circumstances, the subject is his/her own control over time. Apart from single-breath nitrogen washout and bronchial hyperreactivity, most tests are offered only by specialised laboratories. Even if the technology becomes more broadly accessible in multicentre studies, it is anticipated technicians should be trained and follow a quality control programme. Finally, little is known about the MIDs for each test of small airway function.

Conclusion

COPD is a complex disease with diverse pathophysiological manifestations that cannot be adequately assessed by simple spirometry. It follows that clinical trials designed to evaluate the efficacy of pharmacotherapy should not exclusively rely on one physiological marker such as FEV_1. Clearly, there is a need for additional physiological markers which reliably correlate with patient-centred outcomes. Measurement of IC has a solid physiological rationale and correlates well with important patient-centred outcomes such as dyspnoea and exercise tolerance. Inclusion of such measurements in future clinical trials is likely to improve our ability to evaluate efficacy of pharmacotherapy, particularly in patients with more advanced COPD. Lung capacity and volume measurements appear to be adequately responsive to treatment when tested in the multicentre setting. However, further research is urgently needed before these can be accepted as valid physiological markers by clinicians and regulatory authorities tasked with evaluation of new treatments in COPD. Since the primary locus of inflammation in COPD is the peripheral airways, it seems reasonable to directly measure the impact of bronchodilator and anti-inflammatory drug interventions on small airway function, particularly in smokers with preserved FEV_1. Several tests of small airway function are available and appear to be both repeatable and responsive to pharmacotherapy, but extensive validation is required before these can be considered as physiological markers in clinical trials.

Support Statement

P. Laveneziana has received grants from the Seventh Framework Programme of the European Union for the training and career development of researchers (Marie Curie; FP7-PEOPLE-2010-RG) and International Re-integration grants (PIRG07-GA-2010-268396-EDPAH).

Statement of Interest

P. Laveneziana has received grants from Pfizer Investigator-Initiated Research (IIR; grant number WS 942458), and travel/accommodation/meeting expenses from GlaxoSmithKline, France.

References

1. Vestbo J, Hurd SS, Agusti AG, *et al.* Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* 2013; 187: 347–365.
2. Deesomchok A, Webb KA, Forkert L, *et al.* Lung hyperinflation and its reversibility in patients with airway obstruction of varying severity. *COPD* 2010; 7: 428–437.

3. Pellegrino R, Viegi G, Brusasco V, *et al.* Interpretative strategies for lung function tests. *Eur Respir J* 2005; 26: 948–968.
4. Pride NB, Macklem PT. Lung mechanics in disease. *In:* Fishman AP, ed. Handbook of Physiology. Part 2: The Respiratory System. Bethesda, American Physiological Society, 1986; pp. 659–692.
5. Belman MJ, Botnick WC, Shin JW. Inhaled bronchodilators reduce dynamic hyperinflation during exercise in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1996; 153: 967–975.
6. O'Donnell DE, Lam M, Webb KA. Measurement of symptoms, lung hyperinflation, and endurance during exercise in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998; 158: 1557–1565.
7. O'Donnell DE, Lam M, Webb KA. Spirometric correlates of improvement in exercise performance after anticholinergic therapy in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1999; 160: 542–549.
8. O'Donnell DE, Fluge T, Gerken F, *et al.* Effects of tiotropium on lung hyperinflation, dyspnoea and exercise tolerance in COPD. *Eur Respir J* 2004; 23: 832–840.
9. Maltais F, Hamilton A, Marciniuk D, *et al.* Improvements in symptom-limited exercise performance over 8 h with once-daily tiotropium in patients with COPD. *Chest* 2005; 128: 1168–1178.
10. Newton MF, O'Donnell DE, Forkert L. Response of lung volumes to inhaled salbutamol in a large population of patients with severe hyperinflation. *Chest* 2002; 121: 1042–1050.
11. Tashkin DP, Celli B, Decramer M, *et al.* Bronchodilator responsiveness in patients with COPD. *Eur Respir J* 2008; 31: 742–750.
12. Tashkin DP, Celli B, Senn S, *et al.* 4-year trial of tiotropium in chronic obstructive pulmonary disease. *N Engl J Med* 2008; 359: 1543–1554.
13. Swanney MP, Jensen RL, Crichton DA, *et al.* FEV6 is an acceptable surrogate for FVC in the spirometric diagnosis of airway obstruction and restriction. *Am J Respir Crit Care Med* 2000; 162: 917–919.
14. Jensen RL, Crapo RO, Enright P. A statistical rationale for the use of forced expired volume in 6 s. *Chest* 2006; 130: 1650–1656.
15. Akpinar-Elci M, Fedan KB, Enright PL. FEV6 as a surrogate for FVC in detecting airways obstruction and restriction in the workplace. *Eur Respir J* 2006; 27: 374–377.
16. Kainu A, Lindqvist A, Sarna S, *et al.* Responses of FEV6, FVC, and FET to inhaled bronchodilator in the adult general population. *Respir Res* 2009; 10: 71.
17. Donohue JF, Fogarty C, Lötvall J, *et al.* Once-daily bronchodilators for chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2010; 182: 155–162.
18. Cazzola M, MacNee W, Martinez FJ, *et al.* Outcomes for COPD pharmacological trials: from lung function to biomarkers. *Eur Respir J* 2008; 31: 416–469.
19. In chronic obstructive pulmonary disease, a combination of ipratropium and albuterol is more effective than either agent alone. An 85-day multicenter trial. COMBIVENT Inhalation Aerosol Study Group. *Chest* 1994; 105: 1411–1419.
20. Ikeda A, Nishimura K, Koyama H, *et al.* Bronchodilating effects of combined therapy with clinical dosages of ipratropium bromide and salbutamol for stable COPD: comparison with ipratropium bromide alone. *Chest* 1995; 107: 401–405.
21. Aaron SD, Vandemheen KL, Fergusson D, *et al.* Tiotropium in combination with placebo, salmeterol, or fluticasone-salmeterol for treatment of chronic obstructive pulmonary disease: a randomized trial. *Ann Intern Med* 2007; 146: 545–555.
22. Tashkin DP, Littner M, Andrews CP, *et al.* Concomitant treatment with nebulized formoterol and tiotropium in subjects with COPD: a placebo-controlled trial. *Respir Med* 2008; 102: 479–487.
23. van Noord JA, Aumann JL, Janssens E, *et al.* Comparison of tiotropium once daily, formoterol twice daily and both combined once daily in patients with COPD. *Eur Respir J* 2005; 26: 214–222.
24. van Noord JA, Aumann JL, Janssens E, *et al.* Effects of tiotropium with and without formoterol on airflow obstruction and resting hyperinflation in patients with COPD. *Chest* 2006; 129: 509–517.
25. Calverley P, Pauwels R, Vestbo J, *et al.* Combined salmeterol and fluticasone in the treatment of chronic obstructive pulmonary disease: a randomised controlled trial. *Lancet* 2003; 361: 449–456.
26. Calverley PM, Anderson JA, Celli B, *et al.* Salmeterol and fluticasone propionate and survival in chronic obstructive pulmonary disease. *N Engl J Med* 2007; 356: 775–789.
27. Calverley PM, Boonsawat W, Cseke Z, *et al.* Maintenance therapy with budesonide and formoterol in chronic obstructive pulmonary disease. *Eur Respir J* 2003; 22: 912–919.
28. Mahler DA, Wire P, Horstman D, *et al.* Effectiveness of fluticasone propionate and salmeterol combination delivered *via* the Diskus device in the treatment of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2002; 166: 1084–1091.
29. Szafranski W, Cukier A, Ramirez A, *et al.* Efficacy and safety of budesonide/formoterol in the management of chronic obstructive pulmonary disease. *Eur Respir J* 2003; 21: 74–81.
30. Wedzicha JA, Calverley PM, Seemungal TA, *et al.* The prevention of chronic obstructive pulmonary disease exacerbations by salmeterol/fluticasone propionate or tiotropium bromide. *Am J Respir Crit Care Med* 2008; 177: 19–26.
31. Perng DW, Wu CC, Su KC, *et al.* Additive benefits of tiotropium in COPD patients treated with long-acting beta agonists and corticosteroids. *Respirology* 2006; 11: 598–602.

32. Singh D, Brooks J, Hagan G, *et al.* Superiority of "triple" therapy with salmeterol/fluticasone propionate and tiotropium bromide *versus* individual components in moderate to severe COPD. *Thorax* 2008; 63: 592–598.
33. Welte T, Miravitlles M, Hernandez P, *et al.* Efficacy and tolerability of budesonide/formoterol added to tiotropium in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2009; 180: 741–750.
34. Anthonisen NR, Connett JE, Kiley JP, *et al.* Effects of smoking intervention and the use of an inhaled anticholinergic bronchodilator on the rate of decline of FEV1. The Lung Health Study. *JAMA* 1994; 272: 1497–1505.
35. Lung Health Study Research Group. Effect of inhaled triamcinolone on the decline in pulmonary function in chronic obstructive pulmonary disease. *N Engl J Med* 2000; 343: 1902–1909.
36. Burge PS, Calverley PM, Jones PW, *et al.* Randomised, double blind, placebo controlled study of fluticasone propionate in patients with moderate to severe chronic obstructive pulmonary disease: the ISOLDE trial. *BMJ* 2000; 320: 1297–1303.
37. Pauwels RA, Lofdahl CG, Laitinen LA, *et al.* Long-term treatment with inhaled budesonide in persons with mild chronic obstructive pulmonary disease who continue smoking. European Respiratory Society Study on Chronic Obstructive Pulmonary Disease. *N Engl J Med* 1999; 340: 1948–1953.
38. Vestbo J, Sorensen T, Lange P, *et al.* Long-term effect of inhaled budesonide in mild and moderate chronic obstructive pulmonary disease: a randomised controlled trial. *Lancet* 1999; 353: 1819–1823.
39. Celli BR, Thomas NE, Anderson JA, *et al.* Effect of pharmacotherapy on rate of decline of lung function in chronic obstructive pulmonary disease: results from the TORCH study. *Am J Respir Crit Care Med* 2008; 178: 332–338.
40. Hajiro T, Nishimura K, Tsukino M, *et al.* Stages of disease severity and factors that affect the health status of patients with chronic obstructive pulmonary disease. *Respir Med* 2000; 94: 841–846.
41. Miller MR, Hankinson J, Brusasco V, *et al.* Standardisation of spirometry. *Eur Respir J* 2005; 26: 319–338.
42. Quanjer PH, Stanojevic S, Cole TJ, *et al.* Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *Eur Respir J* 2012; 40: 1324–1343.
43. Buist AS, Ross BB. Quantitative analysis of the alveolar plateau in the diagnosis of early airway obstruction. *Am Rev Respir Dis* 1973; 108: 1078–1087.
44. Hogg JC, Chu F, Utokaparch S, *et al.* The nature of small-airway obstruction in chronic obstructive pulmonary disease. *N Engl J Med* 2004; 350: 2645–2653.
45. O'Donnell DE, Forkert L, Webb KA. Evaluation of bronchodilator responses in patients with "irreversible" emphysema. *Eur Respir J* 2001; 18: 914–920.
46. Laveneziana P, Guenette JA, Webb KA, *et al.* New physiological insights into dyspnea and exercise intolerance in chronic obstructive pulmonary disease patients. *Expert Rev Respir Med* 2012; 6: 651–662.
47. O'Donnell DE, Guenette JA, Maltais F, *et al.* Decline of resting inspiratory capacity in COPD: the impact on breathing pattern, dyspnea, and ventilatory capacity during exercise. *Chest* 2012; 141: 753–762.
48. O'Donnell DE, Hamilton AL, Webb KA. Sensory-mechanical relationships during high-intensity, constant-work-rate exercise in COPD. *J Appl Physiol* 2006; 101: 1025–1035.
49. O'Donnell DE, Voduc N, Fitzpatrick M, *et al.* Effect of salmeterol on the ventilatory response to exercise in chronic obstructive pulmonary disease. *Eur Respir J* 2004; 24: 86–94.
50. O'Donnell DE, Laveneziana P. Lung hyperinflation in COPD: the impact of pharmacotherapy. *Eur Respir Rev* 2006; 15: 85–89.
51. Celli B, ZuWallack R, Wang S, *et al.* Improvement in resting inspiratory capacity and hyperinflation with tiotropium in COPD patients with increased static lung volumes. *Chest* 2003; 124: 1743–1748.
52. Chrystyn H, Mulley BA, Peake MD. Dose response relation to oral theophylline in severe chronic obstructive airways disease. *BMJ* 1988; 297: 1506–1510.
53. Guenette JA, Webb KA, O'Donnell DE. Effect of fluticasone/salmeterol combination on dyspnea and respiratory mechanics in mild-to-moderate COPD. *Respir Med* 2013; 107: 708–716.
54. O'Donnell DE, Laveneziana P, Ora J, *et al.* Evaluation of acute bronchodilator reversibility in patients with symptoms of GOLD stage I COPD. *Thorax* 2009; 64: 216–223.
55. O'Donnell DE, Sciurba F, Celli B, *et al.* Effect of fluticasone propionate/salmeterol on lung hyperinflation and exercise endurance in COPD. *Chest* 2006; 130: 647–656.
56. Peters MM, Webb KA, O'Donnell DE. Combined physiological effects of bronchodilators and hyperoxia on exertional dyspnoea in normoxic COPD. *Thorax* 2006; 61: 559–567.
57. Tantucci C, Duguet A, Similowski T, *et al.* Effect of salbutamol on dynamic hyperinflation in chronic obstructive pulmonary disease patients. *Eur Respir J* 1998; 12: 799–804.
58. Laveneziana P, Palange P, Ora J, *et al.* Bronchodilator effect on ventilatory, pulmonary gas exchange, and heart rate kinetics during high-intensity exercise in COPD. *Eur J Appl Physiol* 2009; 107: 633–643.
59. Maltais F, Celli B, Casaburi R, *et al.* Aclidinium bromide improves exercise endurance and lung hyperinflation in patients with moderate to severe COPD. *Respir Med* 2011; 105: 580–587.
60. Beeh KM, Singh D, Di Scala L, *et al.* Once-daily NVA237 improves exercise tolerance from the first dose in patients with COPD: the GLOW3 trial. *Int J Chron Obstruct Pulmon Dis* 2012; 7: 503–513.

61. Berton DC, Reis M, Siqueira AC, *et al.* Effects of tiotropium and formoterol on dynamic hyperinflation and exercise endurance in COPD. *Respir Med* 2010; 104: 1288–1296.
62. Neder JA, Fuld JP, Overend T, *et al.* Effects of formoterol on exercise tolerance in severely disabled patients with COPD. *Respir Med* 2007; 101: 2056–2064.
63. Beeh KM, Wagner F, Khindri S, *et al.* Effect of indacaterol on dynamic lung hyperinflation and breathlessness in hyperinflated patients with COPD. *COPD* 2011; 8: 340–345.
64. O'Donnell DE, Casaburi R, Vincken W, *et al.* Effect of indacaterol on exercise endurance and lung hyperinflation in COPD. *Respir Med* 2011; 105: 1030–1036.
65. Guenette JA, Raghavan N, Harris-McAllister V, *et al.* Effect of adjunct fluticasone propionate on airway physiology during rest and exercise in COPD. *Respir Med* 2011; 105: 1836–1845.
66. Worth H, Forster K, Eriksson G, *et al.* Budesonide added to formoterol contributes to improved exercise tolerance in patients with COPD. *Respir Med* 2010; 104: 1450–1459.
67. Laforce C, Aumann J, de Teresa Parreno L, *et al.* Sustained 24-hour efficacy of once daily indacaterol (300 μg) in patients with chronic obstructive pulmonary disease: a randomized, crossover study. *Pulm Pharmacol Ther* 2011; 24: 162–168.
68. Brusasco V, Canonica GW, Dal Negro R, *et al.* Formoterol by pressurized metered-dose aerosol or dry powder on airway obstruction and lung hyperinflation in partially reversible COPD. *J Aerosol Med Pulm Drug Deliv* 2011; 24: 235–243.
69. Hanania NA, Donohue JF, Nelson H, *et al.* The safety and efficacy of arformoterol and formoterol in COPD. *COPD* 2010; 7: 17–31.
70. Stockley RA, Chopra N, Rice L. Addition of salmeterol to existing treatment in patients with COPD: a 12 month study. *Thorax* 2006; 61: 122–128.
71. Brashier B, Dhembare P, Jantikar A, *et al.* Tiotropium administered by a pressurized metered dose inhaler (pMDI) and spacer produces a similar bronchodilator response as that administered by a Rotahaler in adult subjects with stable moderate-to-severe COPD. *Respir Med* 2007; 101: 2464–2471.
72. Stepanian IE, Khmel'kova NG, Belin-Atarac B. [The efficacy of tiotropium bromide (Spiriva) in the treatment of patients with chronic obstructive pulmonary disease of varying severity: results of the Russian trial]. *Terapevticheskii Arkhiv* 2010; 82: 46–51.
73. Verkindre C, Bart F, Aguilaniu B, *et al.* The effect of tiotropium on hyperinflation and exercise capacity in chronic obstructive pulmonary disease. *Respiration* 2006; 73: 420–427.
74. van der Vaart H, Postma DS, Grevink R, *et al.* Bronchodilation improves endurance but not muscular efficiency in chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis* 2011; 6: 229–235.
75. Fujimoto K, Kitaguchi Y, Kanda S, *et al.* Comparison of efficacy of long-acting bronchodilators in emphysema dominant and emphysema nondominant chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis* 2011; 6: 219–227.
76. Mahler DA, D'Urzo A, Bateman ED, *et al.* Concurrent use of indacaterol plus tiotropium in patients with COPD provides superior bronchodilation compared with tiotropium alone: a randomised, double-blind comparison. *Thorax* 2012; 67: 781–788.
77. Hanania NA, Boota A, Kerwin E, *et al.* Efficacy and safety of nebulized formoterol as add-on therapy in COPD patients receiving maintenance tiotropium bromide: results from a 6-week, randomized, placebo-controlled, clinical trial. *Drugs* 2009; 69: 1205–1216.
78. Tashkin DP, Donohue JF, Mahler DA, *et al.* Effects of arformoterol twice daily, tiotropium once daily, and their combination in patients with COPD. *Respir Med* 2009; 103: 516–524.
79. Magnussen H, Paggiaro P, Schmidt H, *et al.* Effect of combination treatment on lung volumes and exercise endurance time in COPD. *Respir Med* 2012; 106: 1413–1420.
80. O'Donnell DE, Deesomchok A, Lam YM, *et al.* Effects of BMI on static lung volumes in patients with airway obstruction. *Chest* 2011; 140: 461–468.
81. Tzani P, Crisafulli E, Nicolini G, *et al.* Effects of beclomethasone/formoterol fixed combination on lung hyperinflation and dyspnea in COPD patients. *Int J Chron Obstruct Pulmon Dis* 2011; 6: 503–509.
82. Lindberg A, Szalai Z, Pullerits T, *et al.* Fast onset of effect of budesonide/formoterol versus salmeterol/fluticasone and salbutamol in patients with chronic obstructive pulmonary disease and reversible airway obstruction. *Respirology* 2007; 12: 732–739.
83. Celli BR, Tashkin DP, Rennard SI, *et al.* Bronchodilator responsiveness and onset of effect with budesonide/formoterol pMDI in COPD. *Respir Med* 2011; 105: 1176–1188.
84. Dreyse J, Silva F, Diaz O, *et al.* [Clinical and functional benefits of adding theophylline to a standard treatment with short acting bronchodilators in patients with COPD]. *Rev Med Chil* 2005; 133: 1211–1219.
85. Waterhouse JC, Pritchard SM, Howard P. Hyperinflation, trapped gas and theophylline in chronic obstructive pulmonary disease. *Monaldi Arch Chest Dis* 1993; 48: 126–129.
86. Murciano D, Aubier M, Lecocguic Y, *et al.* Effects of theophylline on diaphragmatic strength and fatigue in patients with chronic obstructive pulmonary disease. *N Engl J Med* 1984; 311: 349–353.
87. Murciano D, Auclair MH, Pariente R, *et al.* A randomized, controlled trial of theophylline in patients with severe chronic obstructive pulmonary disease. *N Eng J Med* 1989; 320: 1521–1525.

88. O'Donnell DE, Bredenbroker D, Brose M, *et al.* Physiological effects of roflumilast at rest and during exercise in COPD. *Eur Respir J* 2012; 39: 1104–1112.
89. Stav D, Raz M. Effect of *N*-acetylcysteine on air trapping in COPD: a randomized placebo-controlled study. *Chest* 2009; 136: 381–386.
90. Chin RC, Guenette JA, Cheng S, *et al.* Does the respiratory system limit exercise in mild chronic obstructive pulmonary disease? *Am J Respir Crit Care Med* 2013; 187: 1315–1323.
91. Ofir D, Laveneziana P, Webb KA, *et al.* Mechanisms of dyspnea during cycle exercise in symptomatic patients with GOLD stage I chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2008; 177: 622–629.
92. Wanger J, Clausen JL, Coates A, *et al.* Standardisation of the measurement of lung volumes. *Eur Respir J* 2005; 26: 511–522.
93. Hankinson JL, Stocks J, Peslin R. Reproducibility of lung volume measurements. *Eur Respir J* 1998; 11: 787–790.
94. Casanova C, Cote C, de Torres JP, *et al.* Inspiratory-to-total lung capacity ratio predicts mortality in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2005; 171: 591–597.
95. Cosio M, Ghezzo H, Hogg JC, *et al.* The relations between structural changes in small airways and pulmonary-function tests. *N Engl J Med* 1978; 298: 1277–1281.
96. Mead J. The lung's "quiet zone". *N Engl J Med* 1970; 282: 1318–1319.
97. Hogg JC, McDonough JE, Suzuki M. Small airway obstruction in COPD: new insights based on micro-CT imaging and MRI imaging. *Chest* 2013; 143: 1436–1443.
98. Macklem PT, Hogg WE, Brunton J. Peripheral airway obstruction and particulate deposition in the lung. *Arch Intern Med* 1973; 131: 93–97.
99. Hogg JC, Macklem PT, Thurlbeck WM. Site and nature of airway obstruction in chronic obstructive lung disease. *N Engl J Med* 1968; 278: 1355–1360.
100. Milic-Emili J. Does mechanical injury of the peripheral airways play a role in the genesis of COPD in smokers? *COPD* 2004; 1: 85–92.
101. Verbanck S. Physiological measurement of the small airways. *Respiration* 2012; 84: 177–188.
102. Cockcroft DW, Berscheid BA. Volume adjustment of maximal midexpiratory flow. Importance of changes in total lung capacity. *Chest* 1980; 78: 595–600.
103. Incalzi RA, Pennazza G, Scarlata S, *et al.* Reproducibility and respiratory function correlates of exhaled breath fingerprint in chronic obstructive pulmonary disease. *PloS One* 2012; 7: e45396.
104. Stewart JI, Criner GJ. The small airways in chronic obstructive pulmonary disease: pathology and effects on disease progression and survival. *Curr Opin Pulmon Med* 2013; 19: 109–115.
105. Burgel PR, Bourdin A, Chanez P, *et al.* Update on the roles of distal airways in COPD. *Eur Respir Rev* 2011; 20: 7–22.
106. Incorvaia C, Riario-Sforza GG, Pravettoni C, *et al.* Impairment of small airways in COPD patients with frequent exacerbations and effects of treatment with tiotropium. *Int J Chron Obstruct Pulmon Dis* 2008; 3: 123–126.
107. Wright JL, Lawson LM, Pare PD, *et al.* The detection of small airways disease. *Am Rev Respir Dis* 1984; 129: 989–994.
108. Hanania NA, Celli BR, Donohue JF, *et al.* Bronchodilator reversibility in COPD. *Chest* 2011; 140: 1055–1063.
109. McFadden ER Jr, Kiker R, Holmes B, *et al.* Small airway disease. An assessment of the tests of peripheral airway function. *Am J Med* 1974; 57: 171–182.
110. Contoli M, Bousquet J, Fabbri LM, *et al.* The small airways and distal lung compartment in asthma and COPD: a time for reappraisal. *Allergy* 2010; 65: 141–151.
111. Burgel PR. The role of small airways in obstructive airway diseases. *Eur Respir Rev* 2011; 20: 23–33.
112. Cohen J, Postma DS, Vink-Klooster K, *et al.* FVC to slow inspiratory vital capacity ratio: a potential marker for small airways obstruction. *Chest* 2007; 132: 1198–1203.
113. Sorkness RL, Bleecker ER, Busse WW, *et al.* Lung function in adults with stable but severe asthma: air trapping and incomplete reversal of obstruction with bronchodilation. *J Appl Physiol* 2008; 104: 394–403.
114. Wagner EM, Bleecker ER, Permutt S, *et al.* Direct assessment of small airways reactivity in human subjects. *Am J Respir Crit Care Med* 1998; 157: 447–452.
115. Calverley PM, Kuna P, Monso E, *et al.* Beclomethasone/formoterol in the management of COPD: a randomised controlled trial. *Respir Med* 2010; 104: 1858–1868.
116. Decramer M, Rutten-van Molken M, Dekhuijzen PN, *et al.* Effects of *N*-acetylcysteine on outcomes in chronic obstructive pulmonary disease (Bronchitis Randomized on NAC Cost-Utility Study, BRONCUS): a randomised placebo-controlled trial. *Lancet* 2005; 365: 1552–1560.
117. Bassiri AG, Girgis RE, Doyle RL, *et al.* Detection of small airway dysfunction using specific airway conductance. *Chest* 1997; 111: 1533–1535.
118. Cohen J, Douma WR, van Ooijen PM, *et al.* Localization and quantification of regional and segmental air trapping in asthma. *J Comp Assist Tomogr* 2008; 32: 562–569.
119. Verbanck S, Schuermans D, Van Muylem A, *et al.* Ventilation distribution during histamine provocation. *J Appl Physiol* 1997; 83: 1907–1916.
120. Segal LN, Goldring RM, Oppenheimer BW, *et al.* Disparity between proximal and distal airway reactivity during methacholine challenge. *COPD* 2011; 8: 145–152.
121. Riess A, Wiggs B, Verburgt L, *et al.* Morphologic determinants of airway responsiveness in chronic smokers. *Am J Respir Crit Care Med* 1996; 154: 1444–1449.

122. Oostveen E, MacLeod D, Lorino H, *et al.* The forced oscillation technique in clinical practice: methodology, recommendations and future developments. *Eur Respir J* 2003; 22: 1026–1041.
123. Kjeldgaard JM, Hyde RW, Speers DM, *et al.* Frequency dependence of total respiratory resistance in early airway disease. *Am Rev Respir Dis* 1976; 114: 501–508.
124. Kolsum U, Borrill Z, Roy K, *et al.* Impulse oscillometry in COPD: identification of measurements related to airway obstruction, airway conductance and lung volumes. *Respir Med* 2009; 103: 136–143.
125. Oppenheimer BW, Goldring RM, Berger KI. Distal airway function assessed by oscillometry at varying respiratory rate: comparison with dynamic compliance. *COPD* 2009; 6: 162–170.
126. Timmins SC, Diba C, Farrow CE, *et al.* The relationship between airflow obstruction, emphysema extent, and small airways function in COPD. *Chest* 2012; 142: 312–319.
127. Borrill ZL, Houghton CM, Woodcock AA, *et al.* Measuring bronchodilation in COPD clinical trials. *Br J Clin Pharmacol* 2005; 59: 379–384.
128. Di Mango AM, Lopes AJ, Jansen JM, *et al.* Changes in respiratory mechanics with increasing degrees of airway obstruction in COPD: detection by forced oscillation technique. *Respir Med* 2006; 100: 399–410.
129. Borrill ZL, Houghton CM, Tal-Singer R, *et al.* The use of plethysmography and oscillometry to compare long-acting bronchodilators in patients with COPD. *Br J Clin Pharmacol* 2008; 65: 244–252.
130. Clement J, Landser FJ, Van de Woestijne KP. Total resistance and reactance in patients with respiratory complaints with and without airways obstruction. *Chest* 1983; 83: 215–220.
131. Brochard L, Pelle G, de Palmas J, *et al.* Density and frequency dependence of resistance in early airway obstruction. *Am Rev Respir Dis* 1987; 135: 579–584.
132. Mori K, Shirai T, Mikamo M, *et al.* Respiratory mechanics measured by forced oscillation technique in combined pulmonary fibrosis and emphysema. *Respir Physiol Neurobiol* 2013; 185: 235–240.
133. Paredi P, Goldman M, Alamen A, *et al.* Comparison of inspiratory and expiratory resistance and reactance in patients with asthma and chronic obstructive pulmonary disease. *Thorax* 2010; 65: 263–267.
134. Dellaca RL, Santus P, Aliverti A, *et al.* Detection of expiratory flow limitation in COPD using the forced oscillation technique. *Eur Respir J* 2004; 23: 232–240.
135. Johnson MK, Birch M, Carter R, *et al.* Measurement of physiological recovery from exacerbation of chronic obstructive pulmonary disease using within-breath forced oscillometry. *Thorax* 2007; 62: 299–306.
136. Haruna A, Oga T, Muro S, *et al.* Relationship between peripheral airway function and patient-reported outcomes in COPD: a cross-sectional study. *BMC pulmonary medicine* 2010; 10: 10.
137. Diba C, King GG, Berend N, *et al.* Improved respiratory system conductance following bronchodilator predicts reduced exertional dyspnoea. *Respir Med* 2011; 105: 1345–1351.
138. Timmins SC, Coatsworth N, Palnitkar G, *et al.* Day-to-day variability of oscillatory impedance and spirometry in asthma and COPD. *Respir Physiol Neurobiol* 2013; 185: 416–424.
139. Timmins SC, Diba C, Thamrin C, *et al.* The feasibility of home monitoring of impedance with the forced oscillation technique in chronic obstructive pulmonary disease subjects. *Physiol Meas* 2013; 34: 67–81.
140. Dellacà RL, Pompilio PP, Walker PP, *et al.* Effect of bronchodilation on expiratory flow limitation and resting lung mechanics in COPD. *Eur Respir J* 2009; 33: 1329–1337.
141. Tse HN, Raiteri L, Wong KY, *et al.* High-dose *N*-acetylcysteine in stable chronic obstructive pulmonary disease: the 1-year, double-blind, randomized, placebo-controlled HIACE Study. *Chest* 2013; 144: 106–118.
142. Rodarte JR, Hyatt RE, Westbrook PR. Determination of lung volume by single- and multiple-breath nitrogen washout. *Am Rev Respir Dis* 1976; 114: 131–136.
143. Horsley A. Lung clearance index in the assessment of airways disease. *Respir Med* 2009; 103: 793–799.
144. Verbanck S, Paiva M. Model simulations of gas mixing and ventilation distribution in the human lung. *J Appl Physiol* 1990; 69: 2269–2279.
145. Verbanck S, Schuermans D, Meysman M, *et al.* Noninvasive assessment of airway alterations in smokers: the small airways revisited. *Am J Respir Crit Care Med* 2004; 170: 414–419.
146. Aurora P, Gustafsson P, Bush A, *et al.* Multiple breath inert gas washout as a measure of ventilation distribution in children with cystic fibrosis. *Thorax* 2004; 59: 1068–1073.
147. Lutchen KR, Habib RH, Dorkin HL, *et al.* Respiratory impedance and multibreath N_2 washout in healthy, asthmatic, and cystic fibrosis subjects. *J Appl Physiol* 1990; 68: 2139–2149.
148. Akamatsu K, Matsunaga K, Sugiura H, *et al.* Improvement of airflow limitation by fluticasone propionate/salmeterol in chronic obstructive pulmonary disease: what is the specific marker? *Front Pharmacol* 2011; 2: 36.

Chapter 12

Imaging

David G. Parr

SUMMARY: Contemporary imaging modalities have been employed as biomarkers in a wide range of respiratory diseases. Quantitative assessment of morphology, function and molecular pathogenetic processes using imaging potentially offers significant advantages over traditional outcome measures, particularly in clinical studies of slowly progressive disorders and rare diseases. The variety and complexity of techniques have developed greatly over the last few decades but translation from "bench to bedside" into accepted outcomes for use in clinical studies has lagged behind. Consequently, the extensive literature on methodological detail is not yet appropriately reflected in tangible clinical productivity. Computed tomography densitometry and morphometry are the most validated imaging biomarkers and, consequently, clinical experience is greatest in these modalities, particularly in the field of obstructive airway disease. Progress in the clinical arena will require consensus on appropriate validation strategies and procedural standardisation, and the involvement of industry, and regulatory and reimbursement agencies should be sought at an early stage of development in order to deliver benefit effectively to patients.

Cardio-Respiratory Division, University Hospital Coventry, Coventry, UK.

Correspondence: D.G. Parr, Cardio-Respiratory Division, University Hospital Coventry, Clifford Bridge Road, Coventry, CV2 2DX, UK. Email: david.parr@uhcw.nhs.uk

Eur Respir Monogr 2013; 62: 146–169.
Copyright ERS 2013.
DOI: 10.1183/1025448x.10045512
Print ISBN: 978-1-84984-044-6
Online ISBN: 978-1-84984-045-3
Print ISSN: 1025-448x
Online ISSN: 2075-6674

Imaging has traditionally been performed in routine clinical practice as an aid to clinical diagnosis and to estimate disease severity through visual interpretation. Systematic visual quantification methods have been developed and validated in some respiratory conditions, but such approaches are subjective, poorly reproducible and time consuming. The development of software for automated analysis has enabled the exploitation of digital imaging technology to facilitate objective quantitation. This has led to rapid expansion of the novel field of "imaging biomarkers", a term defined as any anatomical, physiological, biochemical or molecular parameter that may be detected by imaging methods for the purpose of identifying and quantifying disease.

Imaging biomarkers have enormous potential as surrogate end-points for the evaluation of novel medical therapies in early drug development and interventional clinical studies. Improved sensitivity and reproducibility relative to traditional end-points allow the use of smaller sample sizes over a compressed time frame, thereby overcoming the difficulties of patient recruitment and prohibitive costs. These advantages are particularly relevant in slowly progressive and rare diseases or in situations where protracted phase III studies might otherwise be impractical, unethical or unaffordable. Imaging biomarkers also allow quantification of specific disease components for the assessment of targeted treatment in a way that nonspecific global measures, such as forced expiratory volume 1 s (FEV_1), cannot.

Despite the extensive literature on a wide variety of potential imaging biomarkers, the extent to which these methodologies have successfully been applied as clinical outcome measures is varied. The aim of this chapter is to describe the application of the most frequently used imaging biomarkers as outcome measures in clinical studies. The greatest body of experience and published evidence has been in the use of quantitative computed tomography (CT) within the fields of chronic obstructive pulmonary disease (COPD), cystic fibrosis (CF) and asthma. This experience represents a paradigm for the application of other imaging techniques to the quantitation of respiratory disease. Consequently, this chapter will focus on these aspects of quantitative imaging, with discussion of key areas of controversy that are also pertinent to quantitative respiratory imaging in general. Further information on quantitative imaging may be obtained from review articles providing coverage of the wider aspects [1–6].

The importance of structure and structure–function relationships in respiratory disease, and in the development of quantitative imaging methodology

The pathological destruction and remodelling that characterise specific lung diseases result in structural changes in the lung, which in turn are responsible for much of the functional impairment experienced by patients. Consequently, the assessment of structure and appreciation of structure–function relationships in the lung have traditionally underpinned the clinical approach to the diagnosis of respiratory disease, as well as our understanding of the impact of disease on patients. Although the diagnostic gold standard is histopathology, tissue sampling requires invasive methods, and so physiological testing has commonly been employed as the practical alternative in routine clinical practice. Notwithstanding the surrogate nature and limitations of these physiological measures, they are commonly seen as the accepted standard against which new methods are compared.

CT and magnetic resonance imaging (MRI) provide images with structural detail that compares to macroscopic histological examination. *In vivo* morphometry using these imaging modalities therefore represents a tool that compares to the gold standard of histopathology, and the close relationship between structure and function gives clinical relevance to these measurements. The principle advantage of imaging over histopathology is that invasive tissue sampling is avoided and repeated measurements are therefore possible. This is clearly essential for monitoring progressive disease and when specific outcomes are needed in interventional studies of therapy for the attenuation or reversal of disease progression. Imaging-based morphometry has the advantage that a therapeutic response can be assessed within specified regions of the lung that would otherwise remain unidentified using global measures of lung function or localised tissue sampling.

Lung volume variability

A significant obstacle to the performance of accurate and precise morphometry is that the lung is a distensible organ, which consists mostly of air when fully inflated. This complicates the performance and interpretation of imaging-based and histological morphometry [7]. Direct comparison between the two methodologies for the purpose of validation is, consequently, hard to achieve. In addition, fixation-related distortion during histological tissue preparation [8] and the motion of respiration during *in vivo* imaging further confound attempts to draw direct comparison.

It is possible to exert some form of control or apply a correction factor to reduce lung volume variability in the performance of parenchymal assessment. ''Gating'' employs spirometry (or an alternative device) to measure lung volume during image acquisition either for control of inspiratory level at the time of imaging [9, 10] or for retrospective selection of imaging data acquired at designated phases of the respiratory cycle [11]. The volumetric accuracy needed for this approach exceeds that of routine physiological measurement of lung volumes and alternative

methods using modelling for volume correction have been proposed [12–15]. These methods reduce variability to a degree that is acceptable for longitudinal studies and have been recommended for densitometric studies of emphysema [16]. Nevertheless, changes in lung volume, secondary to altered compliance, are also associated with the pathological progression of some conditions; the complexity of these effects may be further compounded by the coexistence of dual pathology [17]. Consequently, the ''correct'' approach to obtaining repeatable measurements in longitudinal studies is contentious.

The influence of inspiratory level on airway dimensions is less predictable than for the parenchyma. In the normal lung, the relative change in airway diameter is linearly related to the cube root of lung volume [18] but this relationship is altered by parenchymal disease. Airway distensibility is reduced by emphysema because airway and parenchymal interdependence is altered by the proteolytic destruction of radial elastin fibres [19, 20]. In addition, the influence of inspiratory level differs according to whether the airways are in the upper or lower lobes [21]. These findings demonstrate the complexity of the performance and interpretation of morphometric studies in the lung but are also of importance in clarifying respiratory pathophysiology.

The assessment of lung parenchyma

Emphysema

COPD is a heterogeneous condition defined by the presence of ''airflow obstruction that is not fully reversible'' and measured in routine clinical practice using the nonspecific spirometric indices of FEV_1 and forced vital capacity. This physiological definition may simplify diagnostic and therapeutic management in routine clinical practice but, arguably, has hampered progress in the development of novel, specific therapies; it is not possible to distinguish the relative contribution of the heterogeneous lung diseases that are united under the umbrella term COPD by this common feature. Furthermore, there is a poor correlation between spirometric indices and symptoms, and poor reproducibility of the measure limits the application of the method as a means of assessing therapeutic efficacy in interventional studies. The need for alternative, more specific, methods for measurement is immediately evident.

CT densitometry

CT imaging enables accurate characterisation of emphysema and visually based severity-scoring systems compare closely with tissue morphometry [22]. In 1984, it was observed that the lung voxel histograms of CT images from patients with emphysema were significantly different from normal (fig. 1) [24]. This led to the use of lung density measurements from CT imaging as a more objective and user-friendly assessment of emphysema *in vivo*. This approach is entirely logical because CT images are densitometric maps, and lung density declines in emphysema due to proteolytic tissue destruction and hyperinflation. A number of densitometric indices derived from the lung voxel histogram have been advocated (fig. 1) [25–28]. The contention that one particular densitometric index is best on the basis of comparison with morbid tissue morphometry is undermined by the inability to demonstrate that *in vivo* dimensions are identical to the *ex vivo* dimensions of fixed tissue. Furthermore, scanner manufacturer, model, settings and protocols are known to influence Hounsfield values and, therefore, apparent voxel density values [16, 29]. A lack of standardisation also prevents comparison between centres and studies. Notwithstanding this caveat, the correlation between CT and pathology is very close for most of the densitometric indices examined [25–28] and, therefore, favouritism for one index over another may be unjustified.

The potential advantages of using CT lung density as a surrogate outcome measure in studies of emphysema-modifying therapy were promptly appreciated and patients with α_1-antitrypsin deficiency (α_1-ATD) were seen to represent a population likely to benefit from such an approach. α_1-ATD-related COPD was considered to be a relatively homogeneous emphysematous phenotype

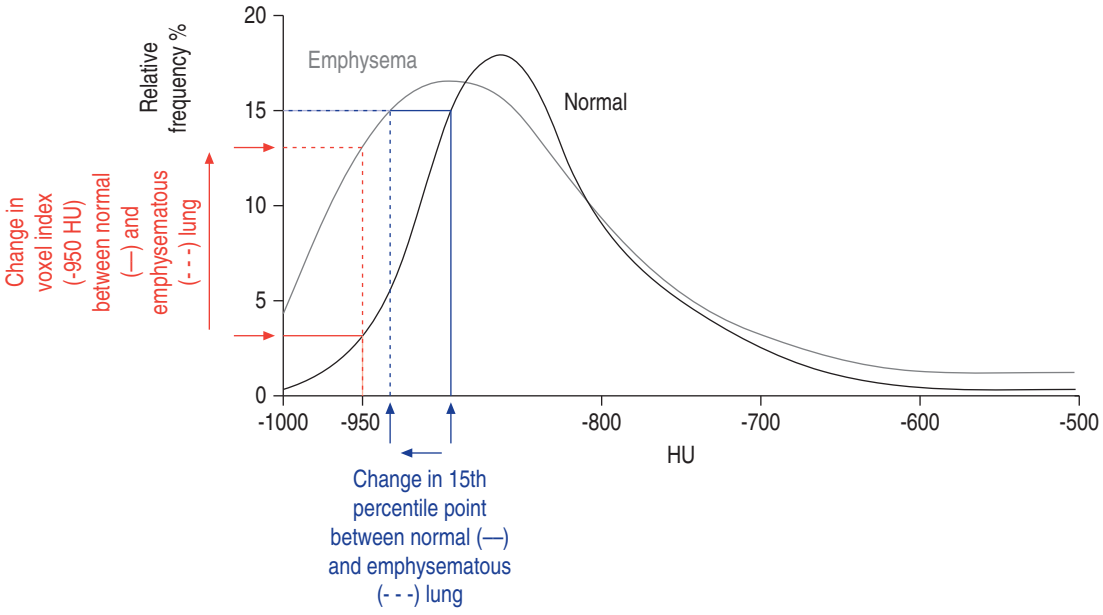


Figure 1. Voxel frequency distribution histogram indicating the appearance in normal lung (black line) and in emphysema (grey line), and the derivation of the commonest densitometric indices. The mean lung density is defined as the mean value (in $g \cdot L^{-1}$) of all voxels distributed within the lung histogram and declines with emphysema progression. The voxel index at a threshold of -950 Hounsfield units (HU), defined as the percentage of voxels with a value less than -950 HU, increases with emphysema progression. The 15th percentile point is defined as the cut-off value, in HU, below which 15% of voxels with the lowest density are distributed. This index becomes more negative with emphysema progression. Reproduced from [23] with permission from the publisher.

and to arise as a result of deficiency of the plasma protein α_1-antitrypsin (α_1-AT). Furthermore, replacement therapy for the arrest or retardation of emphysema progression already existed. However, as a rare lung disease, the practical difficulties of performing a randomised, placebo-controlled trial using FEV1 as the outcome measure were prohibitive [30].

A pilot study of the use of CT densitometry suggested that a study of 130 patients over 3 years would be suitably powered to demonstrate treatment efficacy if CT were the outcome [31]: this compared with estimates of up to 632 patients per arm over 3 years and up to 449 per arm over 5 years if efficacy was assessed using FEV1 [30]. A second "exploratory" randomised placebo-controlled study of α_1-ATD augmentation therapy, EXACTLE (Exacerbations and CT Scan as Lung Endpoints) [32], generated similar results that were suggestive of a beneficial treatment effect on the rate of lung density decline (fig. 2). When treatment efficacy was assessed using targeted regional analysis of the lower lung, which is the predominant location of panlobular emphysema, the retardation of decline in lung density seen in the active arm was statistically significant [23]. Most recently, the results of a suitably powered study (n=180 over 2 years) have demonstrated that intravenous augmentation therapy with an α_1-proteinase inhibitor is effective in retarding the decline in CT-measured lung density. The imaging protocol used percentile density and statistical volume correction, as recommended by an expert working group [16], to demonstrate a reduction in the annual rate of lung density loss in the actively treated arm (-1.45 ± 0.24 *versus* -2.19 ± 0.25 HU; p=0.017) [33].

CT lung densitometry has also been used in a proof-of-concept study (REPAIR) of the use of an oral γ-selective retinoid agonist in α_1-ATD [34]. The treatment did not have a significant benefit on CT lung density but, incidentally, interesting trends were observed in total lung capacity (TLC) and residual volume. The use of a treatment to promote structural repair, rather than to retard emphysema progression (as in α_1-AT augmentation therapy), may influence the choice of

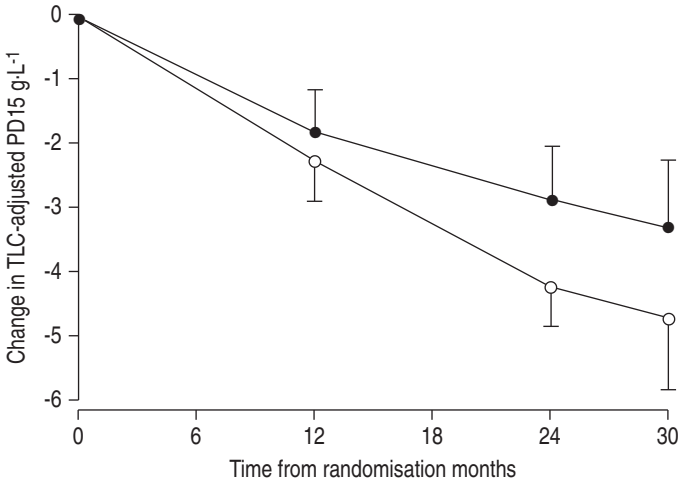


Figure 2. Change from baseline in 15th percentile lung density (PD15), standardised by adjustment to total lung capacity (TLC) in order to correct for differences in inspiratory level between successive scans, over the course of the EXACTLE (Exacerbations and CT Scan as Lung Endpoints) study. Filled circles: active arm (*i.v.* antitrypsin augmentation therapy); open circles: placebo. The error bars indicate standard error of the mean. Estimated treatment difference in mean slopes (annual change in lung density) 0.857 (p=0.068). Reproduced from [32].

methodological approach to reduce volume-related variability. Controlling for differences in inspiration level on sequential imaging may mask important pathophysiological changes: for example, density loss may occur with emphysema progression even in the presence of preserved lung mass, because of worsening hyperinflation. In contrast, therapy-related repair may reverse density loss by reducing compliance and hyperinflation, without necessarily altering tissue mass. This issue is not just of methodological concern, since the conventional understanding that emphysematous damage reduces lung mass has been questioned. On the basis of CT measurements, loss of lung density is still to be observed even after the lung volume changes from progressive hyperinflation are eliminated (fig. 3) [32, 35] and, therefore, lung mass must reduce with emphysema progression. Consequently, where retardation of emphysema progression is the anticipated effect of antiproteolytic treatment, the optimal approach to reducing volume-related variability would appear to be the use of a statistical volume correction method to adjust lung density, measured from imaging acquired at full inspiration and expressed as a percentile index (as recommended by an expert working group [16] and currently employed in antitrypsin augmentation studies). It is less clear whether this is also the best method to use in studies where therapy could alter lung compliance; this is arguably a treatment signal suggested by the physiological data in the REPAIR study [34] but the principle is equally applicable to other chronic lung diseases that are associated with changes in compliance, such as idiopathic pulmonary fibrosis.

CT-derived lung volume measurements

Measurements of lung volume from volumetric CT imaging have consistently been shown to be reliable and comparable to physiological methods, and constitute suitable outcome measures in their own right; the advantage of CT lies in the ability to perform regional assessment. Using this approach, it has been shown that, in the absence of any alteration in total lung volume, the reduction in regional lung volume achieved following placement of bronchial valves correlates with improved symptoms [36]. The localised reduction in lung volume is accompanied by coincident reduction in lung mass, suggestive of reduced perfusion, and compensatory volume changes in the untreated lung. It is unclear which of the observed physiological changes accounts for the symptomatic improvement but the findings have been reproduced in a larger study [37].

CT textural assessment

Attempts have been made to establish techniques that are more specific for emphysema than densitometry. These so-called textural analytical methods use a range of mathematical formulae to describe the grey scale heterogeneity of the parenchyma based on features in the secondary pulmonary lobule [38–42]. They have the potential to distinguish abnormal from normal lung with greater sensitivity than densitometry, and to quantify emphysema subtype (centrilobular, panlobular, *etc.*) [43] and distribution, particularly in the presence of mixed pathology [17].

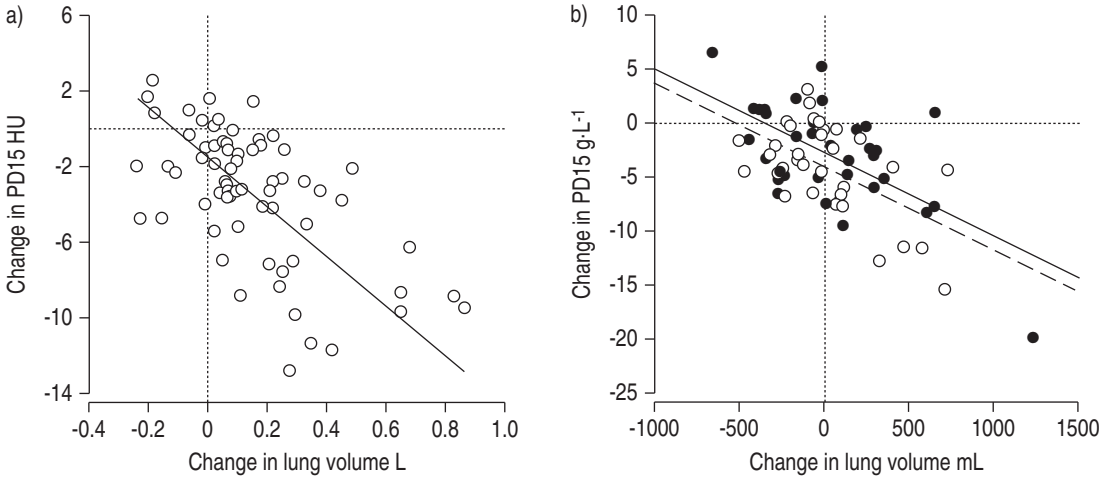


Figure 3. a) Correlation and regression of annual change in 15th percentile lung density (PD15) with annual change in lung volume achieved during image acquisition in an observational study of disease progression in patients with α_1-antitrypsin deficiency measured using computed tomography densitometry. The negative intersect value at $x=0$ is consistent with the loss of lung mass over time. HU: Hounsfield units. Reproduced from [35] with permission from the publisher. b) Plot and regression line of change from baseline to end-point in statistically adjusted PD15 *versus* lung volume. The difference in intersect value at $x=0$ between the active and placebo arms is suggestive of a beneficial treatment effect on the loss of lung mass over time. Filled circles and solid lines: active arm; open circles and dashed lines: placebo. Reproduced from [31] with permission from the publisher.

These methods may represent suitable biomarkers but their role as an outcome measure in interventional studies has yet to be systematically assessed.

Magnetic resonance imaging

MRI has several advantages over CT, including lack of ionising radiation, speed of imaging and the ability to identify tissue characteristics. Conventional MRI is dependent on the presence of protons in the tissue of interest, which become aligned in a strong magnetic field then perturbed by the application of a short burst of radiofrequency energy. On realignment of the protons, energy is released that is detected and converted into an image. The lung has low proton density, which hinders contrast resolution, but the use of contrast agents facilitates assessment of lung function and microstructure.

Helium-3 and xenon-129 may be used as inhaled contrast agents but the technique requires that the gas is hyperpolarised, usually with an onsite laser, shortly prior to a single inhalation into the airways. The sensitivity of static and dynamic imaging of ventilation distribution enables the identification of abnormal heterogeneity even in the presence of normal pulmonary function tests in smoking subjects [44]. In patients with emphysema, the severity of ventilation defects correlates with pulmonary function abnormalities [45] and a quantitative index of the ventilatory defect volume (VDV) (also expressed as a percentage of ventilated lung (VDP)) has been shown to be reproducible [44]. More recently, semiautomated methods for quantifying the ventilated lung have been used [46, 47] but these require further development and validation.

Diffusion imaging uses the high diffusivity of helium-3 or xenon-129 and the application of several radiofrequency pulses to determine the distance travelled by the atoms due to Brownian motion within the "diffusion hindrance" imposed by surrounding lung structures. The effective diffusion distance of helium-3 is limited by the confines of airway and alveolar walls, and this so-called apparent diffusion coefficient (ADC) can provide measurements that correlate closely with histopathology [48]; in contrast, xenon-129 is lipid soluble and diffuses across the alveolar membrane, thereby enabling assessment of both ventilation and perfusion. The methodology has been validated against histopathology in human [48] and animal studies [49], and against the surrogate measures of CT densitometry and gas transfer [50]. In a study of a group of ex-smokers

(with normal CT and spirometry but reduced gas transfer) and patients with COPD (Global Initiative for Chronic Obstructive Lung Disease stage I), ADC measurements correlated with symptoms and exercise capacity [51]. VDV, VDP and ADC measurements are highly reproducible [52] and sensitive to change; a comparison of 15 ex-smokers with COPD and five healthy volunteers over approximately 2 years identified progressive deterioration of MRI indices in the patient group in the absence of significant change in FEV_1 (fig. 4). The authors estimated that 1000 patients and 1000 healthy volunteers would be required to power a study in a comparable patient group with the aim of detecting significant change in FEV_1 [53].

These studies are important in the understanding of the natural history of smoking-related lung disease but also in validation of the methodology. However, the technique is limited by the difficulties in obtaining measurements from poorly ventilated lung and, consequently, becomes less informative with increasing disease severity. It is also unclear whether ADC measurements

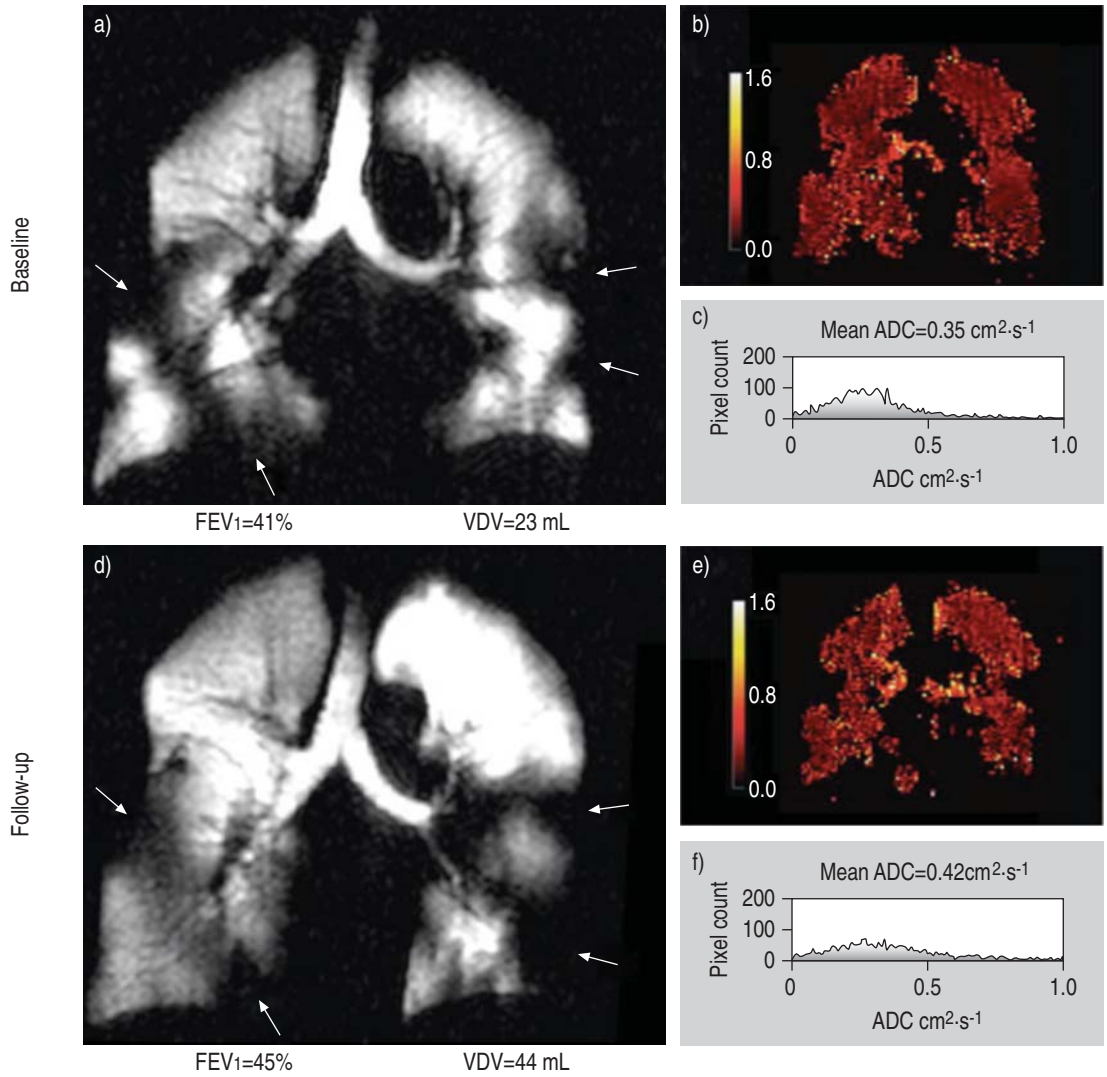


Figure 4. Hyperpolarised helium-3 magnetic resonance data showing typical ventilation defect volume (VDV) and apparent diffusion coefficient (ADC) at a–c) baseline and d–f) follow-up, approximately 2 years later, in a 75-year-old male with stage III chronic obstructive pulmonary disease. a, d) Ventilation images (ventilation defects are indicated by the white arrows). b, e) ADC map. c, f) ADC histogram. VDV and mean ADC increase over time, consistent with disease progression, despite apparent stability in forced expiratory volume in 1 s (FEV_1). Reproduced from [53] with permission from the publisher.

reflect terminal bronchiolar or alveolar dimensions [54, 55]. Attempts have been made to identify regional bronchodilator effects [56] but the interpretation of therapy-associated changes is limited by the lack of suitable clinically meaningful reference standards.

The view that CT is an "antiquated" technique when compared with MRI has to be tempered with some practical considerations. There is concern over the limited availability and expense of helium and, although most hyperpolarised helium-3 MRI techniques are directly translatable to hyperpolarise xenon-129, the anaesthetic properties of this gas need to be considered. Notwithstanding this concern, more recently, oxygen-enhanced MRI techniques have been used to assess lung ventilation [57] and these may obviate the need for hyperpolarised noble gas contrast agents. The development of quantitative assessment tools in MRI is problematic; signal intensities are not calibrated as compared with CT, and the signal-to-noise ratio in lung tissue is low and heterogeneous due to several physical circumstances [58]. Furthermore, the hyperpolarised MRI community has not yet established standard imaging techniques.

The assessment of airway disease

Chronic obstructive pulmonary disease

CT airway morphometry

Quantitative characterisation of airway disease is less well established as an outcome in interventional studies than techniques for parenchymal assessment. CT is currently the method of choice for airway morphometry even though the principal site of expiratory airflow obstruction lies beyond the resolution of clinical scanners, in airways with a diameter below 2 mm. Despite this limitation, it is known that the central airways reflect remodelling in the lung periphery [59], as borne out by the finding that wall thickening of the central airways relates to the severity of small airway disease [60]. A number of different approaches have been implemented in automated algorithms to quantify the dimensional changes that reflect remodelling, wall thickening and luminal narrowing secondary to inflammation. One of the most commonly used metrics is the cross-sectional area of the airway wall, expressed as a percentage of the total cross-sectional area of the airway (WA%) (fig. 5), primarily because it has been consistently shown to relate to airflow obstruction [61–63]. However, WA% tends to increase as total bronchial area decreases, so that a standardised metric is required. Since the square root of the airway wall area is linearly related to the internal perimeter (P_i) of the airway [60], the square root of the cross-sectional area of the airway wall for an arbitrary airway with a P_i of 10 mm can be predicted by linear regression. This standardised metric can be used to compare measurements between subjects and is effective in phenotyping studies. As with densitometry, airway morphometry is influenced by scanner type, imaging protocol and post-imaging processes [64, 65] and the optimum methodology remains contentious.

The development of multidetector CT and multiplanar reconstruction techniques enables volumetric analysis of the entire bronchial tree so that accurate analysis of the more peripheral airways is now considered possible (fig. 6) [62]. Notwithstanding this belief, it is uncertain whether accurate measures really can be obtained beyond the third- or fourth-generation subsegmental airways, and what the changes to airway dimensions represent. Correlations between airway metrics and lung function are better for more peripheral airways [62] but this relationship may actually be a reflection of the influence of juxtaposed emphysema on WA% measurements of small airways, rather than of intrinsic airways disease [66]. Additional credence to this explanation comes from the finding that the value of CT measurements of airway wall thickness in the prediction of mortality is dependent on emphysema severity [67]. These data further demonstrate the need to improve techniques for airway metrics in order to disentangle the apparent interdependence between CT-based airway morphometry and emphysema quantitation. The extrapolation of central airway dimensions in assessing small airway disease is also brought into question by data from micro-CT examination of resected lung [68] and from clinical imaging in smokers [69], which indicate that airflow obstruction arises from outright loss of airways and not

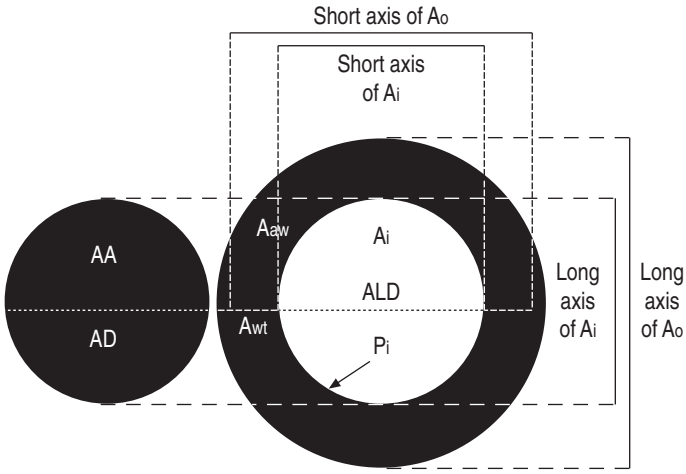


Figure 5. Computed tomography-estimated dimensions of the airway lumen and wall, and accompanying pulmonary artery. Linear measurements include: the airway perimeter (P_i); the long and short axes of the outer airway area (A_o) and the lumen area (A_i); wall thickness (A_{wt}); airway lumen diameter (ALD); airway outer diameter (ALD+Awt × 2); and arterial diameter (AD). The area dimensions include: A_i; airway wall area (A_{aw}); outer airway area ($A_o = A_{aw} + A_i$); and arterial area (AA). Percentage wall area is derived from $A_{aw}/A_o \times 100\%$. Reproduced from [61].

from airway remodelling alone. Consequently, airway wall thickening may reflect disease severity rather than represent a meaningful therapeutic target. Notwithstanding these fundamental concerns, the validity of this biomarker is supported by the utility of CT-based quantification methods in demonstrating that the parenchymal and airway components of COPD contribute independently to airflow obstruction [70] and have distinct heritability [71]. Airway morphometry has also been validated against a pathological standard using micro-CT [72] and clinical scanners [73–75]. However, this approach is subject to the same histopathological limitations as described above for emphysema studies and the use of airway phantoms is limited by their poor representation of *in vivo* airways.

CT airway densitometry

Smoking-related changes also appear to influence airway wall density, since CT-derived measurements of peak mural density correlate with airflow obstruction (more strongly with sixth-order than with third-order bronchi) [76]. It has been suggested that mural density or attenuation reflects alterations in the composition of the airway wall [77] but, in the absence of histopathological correlation, the use of this metric as an outcome measure is currently unvalidated.

Air trapping

Indirect approaches have been used to assess the small airways by the measurement of "air trapping" through densitometry of expiratory imaging; estimation of the appropriate voxel index threshold is based on the theoretical X-ray attenuation value that represents maximal lung inflation, *i.e.* TLC divided by lung mass. The application of a threshold of -856 HU may identify air trapping in subjects with asthma [78], but in COPD it is necessary to pair inspiratory and expiratory data in order to distinguish emphysema-related gas trapping from that due to small airway disease [79]. Whether this technique, when sufficiently validated, justifies the use of ionising radiation to replace a physiological measure remains to be seen.

Asthma

Asthma is defined by airway hyperresponsiveness with episodic airflow limitation, caused by airway inflammation and remodelling, and associated with detectable morphological changes in the airways and parenchyma. Conventionally, clinical characterisation is based on measurement of airflow obstruction and the effectiveness of treatment is assessed by control of symptoms and airflow obstruction. Imaging offers a noninvasive method for quantifying airway dimensions but can also provide functional metrics with a spatial component.

CT morphometry

Airway remodelling is considered to have an important pathophysiological role in asthma [80] and, consequently, imaging studies have sought to explore this feature as a potential therapeutic outcome measure. Wall thickening in the central airways worsens with increasing asthma severity [81]

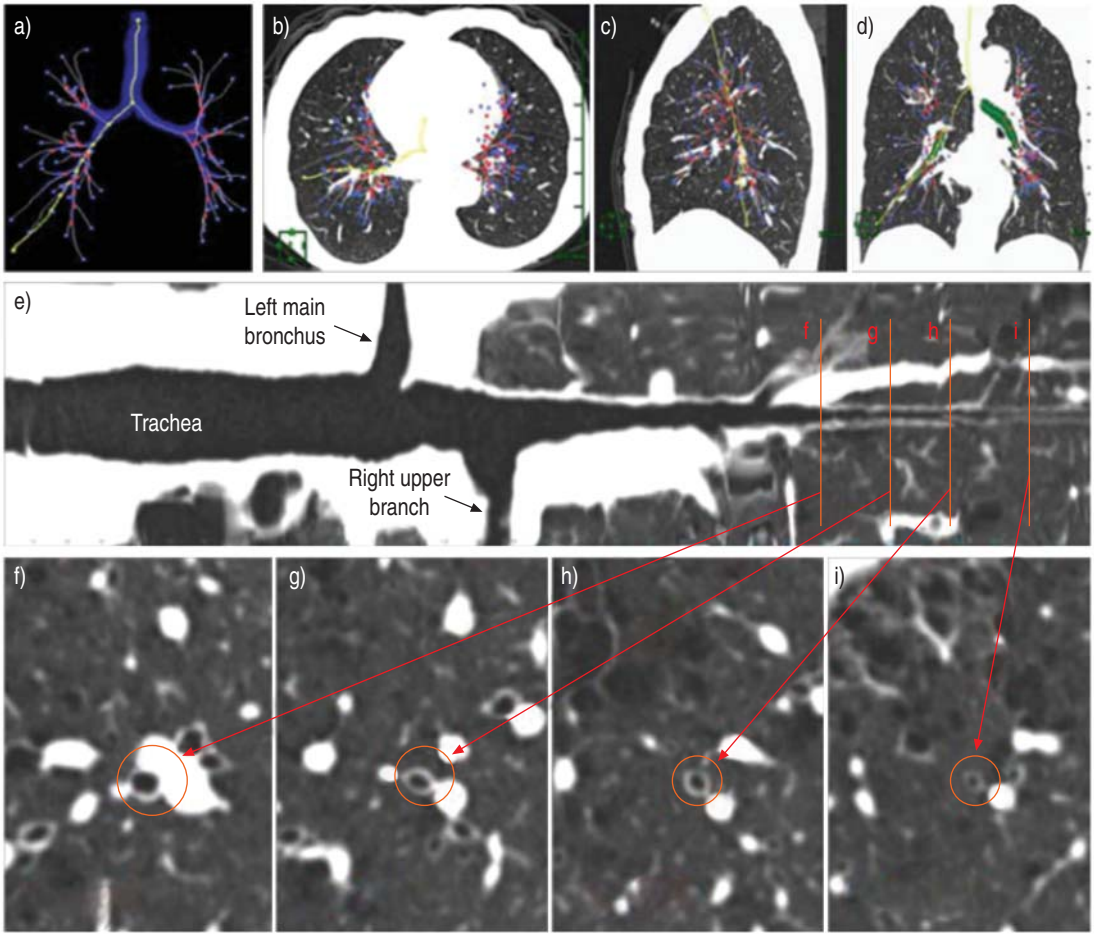


Figure 6. Sequential image processing procedure for multiplanar reconstruction (MPR) to obtain short-axis images that are exactly perpendicular to the long axis in any location. a) Construction of an accurate bronchial skeleton. Source images of b) axial, c) sagittal and d) coronal slices. Yellow lines indicate the same series of airway in a–d. e) A curved MPR image of the selected airway. f–i) Short-axis images obtained from the curved MPR image and oriented precisely perpendicular to the long axis of the airway. The numbers of the bronchial generation are shown. Reproduced from [62] with permission from the publisher.

and is positively correlated with pulmonary function in the stable state [81, 82]. Case reports would also suggest that airway dimensions are altered acutely and reversibly at the time of an exacerbation. In patients with chronic cough, airway wall thickening is seen in both cough-variant asthma and in subjects with nonasthmatic cough [83]. As intimated by the authors of this study, although the magnitude of thickening was greater in the asthmatic group, the technique is essentially a nonspecific index of the common effect of differing pathologies. Notwithstanding this limitation, changes in central airway dimensions can be identified in response to therapy and, in carefully selected phenotypes, may be used as an informative outcome. Measurements of wall thickening and luminal area from the right upper lobe apical segment bronchus suggest beneficial treatment effects of an anti-IgE antibody [84], and of inhaled corticosteroid and long-acting β_2-agonist in combination [85]. In both studies, these changes were associated with improvements in FEV_1, sputum eosinophil counts and health status.

Air trapping

As with COPD, when comparison is made between the central and distal airways, the correlation with airflow limitation is stronger for the third- to fifth-generation bronchi [86]. As a consequence of increasing interest in the smaller airways tempered by the limits of CT resolution, some

investigators have pursued densitometric measurements of expiratory scans as an indirect measure of small airway function. Air trapping on expiration may be quantified using a voxel index threshold of -856 HU, as described previously for COPD. Such measurements correlate with clinical symptoms [78] and change in association with therapy [87]. They may also discriminate between different therapies where global physiological measures of small airways function remain unchanged [88]. Whether this indicates insensitivity of physiological tests or a lack of clinical significance of the CT changes is undetermined.

Magnetic resonance imaging

Contrast-enhanced MRI has been used to explore small airway function in asthma, in particular, to evaluate regional responses to intervention. Ventilation defects on hyperpolarised helium-3 MRI imaging correlate with asthma severity and spirometric measures of airflow obstruction [89]. The number and size of such ventilation defects increase following methacholine challenge, in parallel with decreases in FEV_1 [90]. When imaged over longer intervals, some defects appear to be transient; however, the majority appear to remain in fixed locations regardless of disease severity [91] and so the assessment of regional responses to bronchodilator therapy may be reliable. Helium-3 and xenon-129 have both been used for such studies, but it has been shown that there are important differences in the sensitivity of the methods [92].

CF and bronchiectasis

The predominant cause of mortality in CF is progressive obstructive lung disease arising from inflammation and tissue destruction. Spirometry is the traditional tool for assessing disease progression and predicting survival, and the most validated surrogate end-point is FEV_1. However, this is a unidimensional measure of a heterogeneous condition.

Computed tomography

CT imaging can delineate the morphological changes of bronchiectasis, bronchial wall thickening, mucus plugging, air trapping, mosaic perfusion and ground-glass opacification that characterise CF-related lung disease. This wide spectrum of morphological change is unevenly distributed throughout the lung and varies according to disease severity [93]; consequently, only crude evaluation can be achieved using a global measure of lung function, such as FEV_1. CT can be performed across all age groups, including infants, and is particularly advantageous for the evaluation of early or mild CF lung disease and severe advanced disease. Cross-sectional studies in infants and young children have demonstrated greater airway wall thickness and lumen diameter compared with matched control subjects [94], and mild bronchiectasis or mosaicism can be detected in patients with a normal FEV_1 [95]. Cohort studies have demonstrated that stability [96] or even improvement [97] in FEV_1 may coexist with progressive morphological change on CT.

The complexity of morphological changes in CF-related lung disease has meant that, with the exception of automated airway morphometry [98] and densitometric assessment of air trapping [99], quantitative CT studies have been largely restricted to the use of visual scoring systems [100–102]. These systems are reproducible [103, 104] and less subject to the confounding variable of inspiratory level than lung densitometry (with the exception of the assessment of air trapping from expiratory images, when gating methods are advocated in young children). Detail relating to the relative progression of specific disease components and spatial distribution can be assessed. Comparative studies have demonstrated that CT is nearly twice as sensitive at detecting change as lung function, with an ability to detect increases in airway wall thickening of 1% per year [105]. The most evident changes of progression are those of bronchiectasis and mucus plugging [97, 105]. Bronchiectasis is known to be largely irreversible and, consequently, constitutes a more readily identifiable outcome measure than potentially reversible changes such as mucus plugging or bronchial wall thickening. Nevertheless, the relevance of specific features may differ with disease stage, *e.g.* air trapping is seen to be a suitable specific therapeutic target that is reversible when treated early.

Parenchymal densitometry, a surrogate for air trapping, has been shown to be more consistent and sensitive than either spirometry or total CT scores. In a small randomised controlled study, treatment with dornase alfa reduced air trapping on CT and improved peripheral airway obstruction in patients with mild to moderately severe CF lung disease [99]. The superior sensitivity to change of CT has also been confirmed in other interventional studies using dornase alfa [99, 102, 106, 107] and conventional treatment for acute exacerbations [108–110]. Consequently, the use of CT enables a reduction in sample size and study duration in interventional studies. It has been estimated that, because of improvements seen in retardation of lung function decline due to contemporary therapies, at least 1000 patients would now be required to run PEIT (Pulmozyme Early Intervention Trial) employing FEV_1 as the end-point, whereas only 150 would be required using CT [111]. The combination of spirometry and CT as a composite outcome measure further increases sensitivity [107], and this end-point may be perceived as having greater clinical relevance because of the inclusion of an index of function.

The validity of CT as an outcome measure in CF is hard to prove in the absence of a gold standard. CT measures of non-CF bronchiectasis relate closely to histopathological changes [112] and inflammation measured from bronchoalveolar lavage samples [113], but similar studies have not been performed in CF-related lung disease. Airway morphometry correlates with lung function [98] and CT scores relate to clinical end-points, such as infective exacerbations [114]. More recently, a study of CT in patients awaiting lung transplantation found that CT scores predicted survival [115]. This latter finding is clearly of importance in routine clinical treatment planning but also validates the methodology against a clinical end-point that is accepted to be ''meaningful''. As with the use of CT in COPD, agreement is still required on a standardised methodology, although semiquantitative scoring systems appear not to be subject to the same variability arising from imaging protocols and inspiratory volume during scanning that influence densitometry. At present, both inspiratory and expiratory scans are considered necessary and the use of volume control has been advocated [116].

It is uncertain whether CT represents a suitable primary outcome measure or an important secondary end-point to clarify mechanism of action or dose response, or to define time course of treatment. While there are clear advantages to CT, these need to be balanced against the risks, which include the use of sedation in children and exposure to radiation. Understanding of the degree of risk arising from radiation is becoming more evidence based, and this is increasingly important given the young age at exposure and the continuing improvements in survival and longevity [117]. Overall cancer incidence rises by 24% with exposure and this increases following exposure at younger ages. Estimates suggest that, for each additional CT scan, the incidence increases by 16% compared with nonexposure [118].

Magnetic resonance imaging

MRI is established as the radiation-free alternative to CT. Common proton MRI sequences can provide data on airway morphology down to the third and fourth generations and on bronchial wall thickening, depending on the bronchial size and signal [119, 120]. Air fluid levels, air bronchograms and parenchymal destruction can usually be differentiated similarly to CT but, in addition, the identification of inflammation-related oedema and the differentiation of mucus plugging from bronchial wall thickening are achievable on T2 weighted images [120]. Semiquantitative visual scoring systems have been adapted from CT methods [119, 121], and appear to be clinically valid (based on good correlation with CT and spirometric indices) [122] and responsive to treatment [123].

Functional MRI techniques have high sensitivity for the detection of perfusion and ventilation abnormalities. Even in the absence of clear morphological changes on MRI, perfusion defects can be identified, possibly due to air trapping [124]. Inhomogeneous lung ventilation [125] and ventilation defects [126] can be identified using oxygen-enhanced MRI and hyperpolarised helium-3 MRI, respectively. The use of semiquantitative scoring of ventilation defects has been explored and shown to correlate with spirometric indices [127]. While these functional imaging

techniques have been of a predominantly descriptive nature, methodological development has, more recently, seen the generation of quantitative data [128].

Interstitial lung disease

Computed tomography

Histogram-based assessment of lung density has been applied to the quantitation of fibrotic lung disease but with less success than in emphysema. This heterogeneous group of conditions has historically been viewed as having an inflammatory cause with a fibrotic end stage that are both reflected in reduced lung compliance, increased tissue mass and increased density. In parallel with more discriminating contemporary views of pathogenesis, there has been a shift away from first-order densitometric measures (*e.g.* mean lung density, skewness and kurtosis) towards more representative methods such as textural analysis [38, 39]. It is now possible using automated software to differentiate between normal lung and specific imaging abnormalities such as honeycombing, emphysema, consolidation, comparable to visual interpretation (table 1 and fig. 7) [129, 130]. Longitudinal data indicate that interval changes in specific features can be reliably quantified and that these contribute differentially to the decline in physiological indices [130].

The structural and physiological complexity of other forms of interstitial lung disease (ILD) have not prevented attempts to develop quantitative imaging techniques. For example, parenchymal textural changes in sarcoidosis [131] and lymphangioleiomyomatosis [132] have been quantified using automated software and correlated with physiological impairment. These preliminary findings suggest a potential role for quantitative imaging in these conditions, although there may be better alternatives to morphological indices (see later).

Molecular imaging methods

Imaging techniques that use labelled markers offer an opportunity to perform repeated noninvasive measurements of specific cell behaviour and biochemical events *in vivo*. The glucose analogue fluorine-18-labelled 2-fluoro-2-deoxy-D-glucose (FDG) is the most common PET tracer and is used in the majority of clinical PET imaging. FDG is taken up by cells in proportion to their metabolic activity then phosphorylated to FDG-6-phosphate. Because glycolytic alteration is not possible, the tracer is "trapped" and subsequent decay of the isotope provides a signal that represents the cell's metabolic activity. FDG PET/CT is now used routinely to stage solid lung tumours and may also indicate therapeutic responsiveness. However, chronic inflammatory lung diseases do not have the same favourable combination of high glycolytic activity plus high cellular density seen in solid tumours. Consequently, the imaging protocols required for quantitative assessment of inflammatory cell activity in these conditions differ considerably from those used in routine oncological practice [133, 134].

Pulmonary uptake of FDG in animal models of lobar pneumonia [135] and fibrosis [136] has been shown to relate to neutrophilic activity. Therefore, FDG PET/CT may be a valid noninvasive clinical biomarker of pulmonary neutrophilic activity in conditions associated with neutrophil

Table 1. Comparison of fractional volume of abnormalities assessed using automated textural analysis and according to a radiologist's interpretation of the images in figure 7

	Fractional volume %							
	NL	GGO	RO	HC	EMPH	CONS	Abnormal	Fibrosis
System	34	11	23	29	3	0	66	52
Reader	30	15	5	45	5	0	70	50

NL: normal lung; GGO: ground-glass opacity; RO: reticular opacity; HC: honeycombing; EMPH: emphysema; CONS: consolidation. Reproduced from [129] with permission from the publisher.

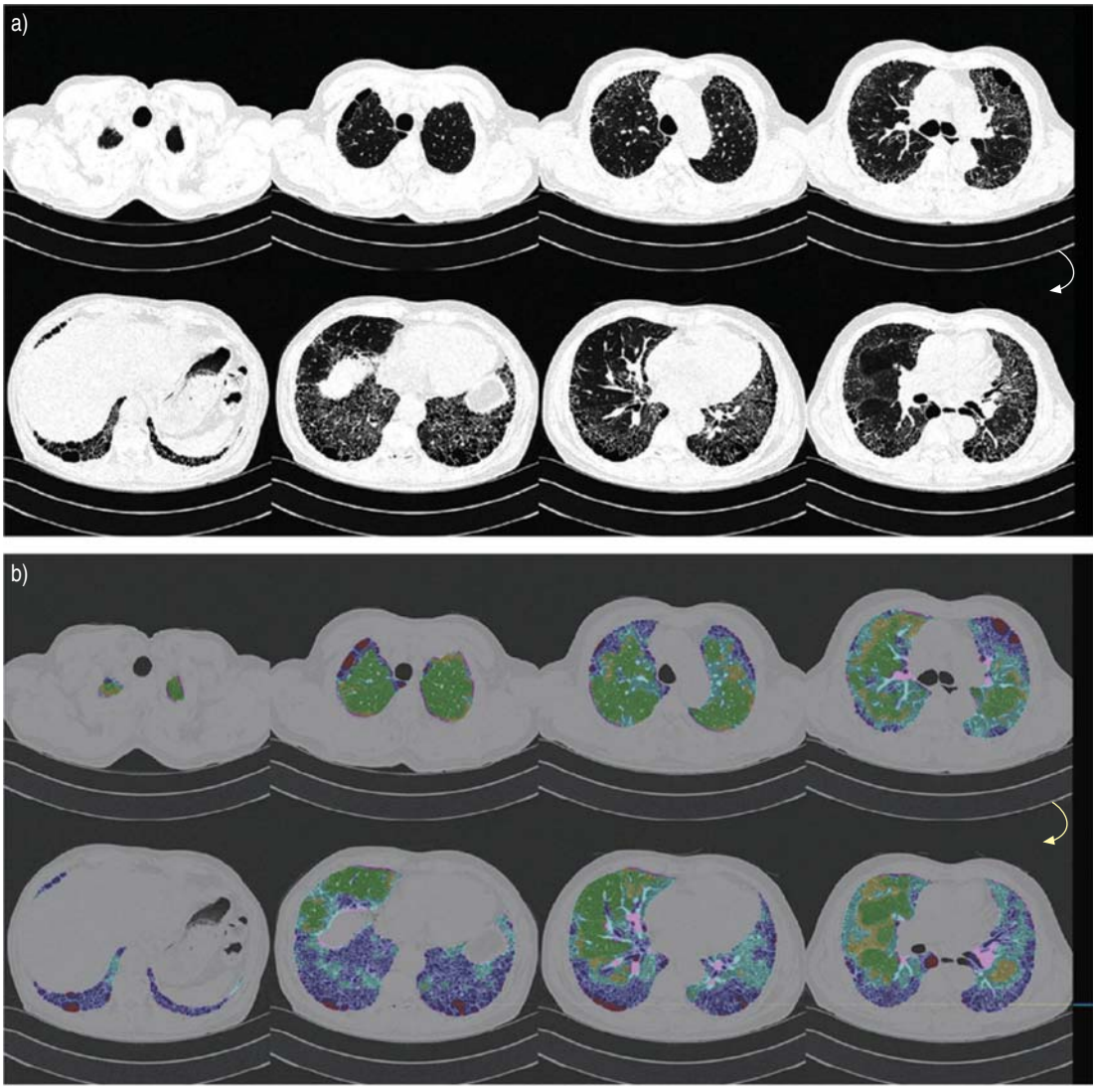


Figure 7. Computed tomography (CT) imaging of a patient with usual interstitial pneumonia. a) CT images predominantly showed basal and peripheral honeycombing (blue) and reticular opacities. b) Textural analysis using an automated system showing approximately one half of the lung volume as honeycombing and reticular opacities. Quantification of these images is presented in table 1. Green: normal lung; yellow: ground-glass opacity; cyan: reticular opacity; red: emphysema; pink: consolidation. Reproduced from [129] with permission from the publisher.

inflammation. This kind of signal is not evident on visual inspection of images, even in the presence of significant inflammation, because of low tissue density and high plasma/tissue ratio of activity. However, a signal is detectable when FDG uptake is quantified in relation to tissue mass rather than volume [133]. This method has been applied to several chronic lung diseases.

Obstructive airway disease

Pulmonary FDG uptake is significantly higher in patients with COPD than in asthma and healthy control subjects, even when studying small group sizes (n=6) [137]. Recent confirmation of this finding in COPD has shown, using three-dimensional assessment of FDG PET/CT imaging, that FDG uptake is greatest in the upper lung in association with apical emphysema (fig. 8) [134]. Correlation of the signal with densitometric and physiological measures of emphysema severity provides further

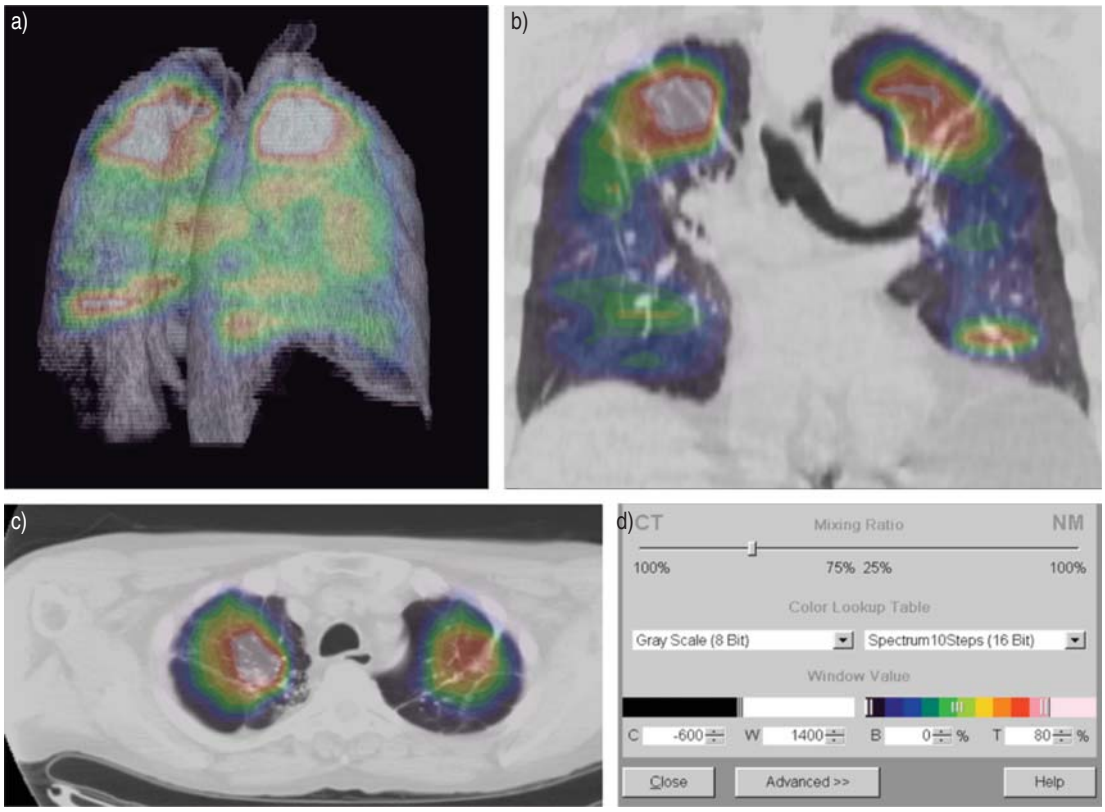


Figure 8. a) Three-dimensional, b) axial and c) coronal imaging illustrating the predominantly apical distribution of pulmonary fluorine-18-labelled 2-fluoro-2-deoxy-D-glucose (FDG) uptake in a patient with usual chronic obstructive pulmonary disease. d) The colour scale gives the spectrum applied to the full range of FDG uptake, estimated by Patlak analysis, so that each colour band represents a range of 10% of the maximum signal, with the maximum signal represented by white and the minimum signal by black. Reproduced from [134] with permission from the publisher.

evidence that FDG uptake in COPD is specifically related to the presence of emphysema. Nevertheless, despite the known association between neutrophilic inflammation and α_1-ATD-related emphysema, there was no difference in uptake between this group and healthy control subjects.

Unexpected data have also been generated in studies of bronchiectasis. The loss of isotope from the body following *i.v.* injection of indium-111-labelled granulocytes is interpreted as reflecting neutrophil migration into the airways, as the signal correlates with sputum volume and the extent of bronchiectasis on CT [138]. However, the high uptake of FDG that might be expected on PET imaging is not seen [133]. In contrast, increased pulmonary FDG uptake is seen in CF-related bronchiectasis and this relates to sputum neutrophil counts, pulmonary function and CT severity scores [139]. FDG uptake is also seen to rise during CF exacerbations and fall following therapy-associated resolution [139, 140], thereby providing a potential biomarker of neutrophilic inflammation in this condition.

In a study of atopic asthma, airway inflammation invoked by the localised instillation of allergen may be visualised using FDG and PET [141]. Whether the tracked granulocytes are eosinophils rather than neutrophils is not clear.

Interstitial lung disease

Preliminary data suggest a role for FDG in the monitoring of sarcoidosis [142, 143] with uptake reflecting changes in serum angiotensin-converting enzyme and predicting functional decline [144].

More detailed studies have been performed in diffuse parenchymal lung disease, including idiopathic pulmonary fibrosis; increased FDG uptake is seen in association with specific CT features, particularly honeycombing, and this relates to health status and physiology [145]. Dual-isotope imaging using a labelled somatostatin receptor analogue and FDG identifies a co-distribution of uptake that relates to physiological measures [146]. Such an approach demonstrates the potential utility of this methodology for use in pre-clinical and early clinical studies of novel therapies in ILD.

While these data support the use of the methodology as a noninvasive tool for clinical studies of anti-inflammatory therapies, the apparent inconsistent findings seen in some conditions require further clarification. In addition, there needs to be improved understanding of how FDG uptake relates to inflammatory cell function.

Future development of quantitative imaging methods

The area of greatest promise for the future is believed, by many in the field, to be that of functional imaging, and this is reflected in the rapidly increasing volume of published studies. Functional imaging has yet to be applied routinely to clinical trials but the use of CT, PET, MRI and techniques such as SPECT (single photon emission computed tomography) offer a potential means of providing novel insight into pulmonary physiology and pathophysiology [1, 4, 147–149]. The application of computational modelling techniques, such as computational fluid dynamics, may represent a means of gaining better understanding of airflow within the airways [147, 150–154]. However, the validation of these ''functional'' methods is problematic; the detection of an imaging signal under circumstances when conventional clinical measures lack sensitivity, an often-cited advantage of these novel techniques, makes validation very difficult if the conventional measure happens to be the gold standard. Interpretation of the clinical relevance of such techniques is also limited in the absence of a clear relationship to a clinical reference. Validation using animal models is an option, provided that such an approach is acceptable to regulatory bodies. However, attempts at validation by comparison against other imaging methods [154] are of questionable value unless these techniques have previously been accepted as valid standards.

Thoracic malignancies have a poor prognosis and response to treatment, so that the most commonly utilised outcome in clinical trials has, historically, been that of mortality. Nevertheless, quantitative CT, MRI and PET imaging are expected to play an increasing role in oncology beyond that of the measurement of tumour volume, with the potential to objectively assess molecularly targeted, personalised therapies [155, 156]. The use of molecular imaging probes, such as radioisotopes in PET or gadolinium in MRI, may represent a noninvasive, surrogate measure for the gold standard, histopathological response. These methods may be used to monitor the response of tumours to therapy in routine clinical practice as well as for end-points in translational pharmacodynamic studies.

The use and acceptance of imaging biomarkers as valid end-points in interventional studies

The development and acceptance of novel methodologies in clinical medicine often requires painstaking validation over many years, and progress occurs at a frustratingly slow rate for proponents and patients alike. For example, the concept of spirometry dates back to the time of the Roman Empire, when Galen performed a volumetric experiment on human ventilation using a bladder, yet the methodology for obtaining forced expiratory manoeuvres and derived parameters was not standardised until 2005 [157]. It is now nearly 30 years since the first imaging biomarker, CT-measured lung density, was proposed as a potential noninvasive surrogate measure of emphysema [24] but the expectation that this technique should have received widespread acceptance may be unrealistic, particularly when the track record of spirometry is considered.

The role of the surrogate end-point is contentious. In phase I and II trials, surrogate end-points are considered acceptable primary end-points by regulatory authorities such as the US Food and Drug Administration (FDA). Consequently, imaging biomarkers may have a role in the early phases of drug development, *e.g.* to select promising new therapies from the large number of potential candidates, or to facilitate the transition from animal model experiments to human studies with dose-ranging studies or validation of presumed therapeutic mechanisms.

The stance of regulatory authorities towards imaging biomarkers in clinical studies is less clear. As with any surrogate end-point, validation requires that the relationship between novel imaging biomarkers and clinical end-points be defined. They need to be biologically plausible, reflect clinical severity and correlate with true outcome measures [158]: they should be responsive, such that the measurement changes in the anticipated direction either with disease progression or with an intervention. The methodologies need to be standardised and the strengths and limitations understood in order to ensure that the trial end-points are appropriate: consideration must be given to the phase of trial, the study population (in terms of disease phenotype and severity), the duration of study, the sample size and the therapeutic aim. In addition, it is desirable that the end-points are applicable across all age groups, and the techniques are easy to perform, of minimal risk and inexpensive. Careful consideration should also be given to the number and expertise of participating sites, particularly when technically demanding methods are to be used.

The expectation of the FDA that the primary end-point in phase III studies must be clinically meaningful has been subject to the influence of a number of changes to statutory regulations. The Food and Drug Modernization Act of 1997 [159] was enacted to address perceived shortcomings in the existing medical product regulatory process, such as the length of time required for the approval of drugs used in the management of serious life-threatening conditions. Furthermore, marketing approval for a drug may be secured when the FDA considers that the product has an effect on a surrogate end-point that is reasonably likely to predict clinical benefit. Under these circumstances, imaging biomarkers can contribute to the requirements of post-market approval surveillance. In the field of rare diseases, the pursuit of the perfect "clinically meaningful" outcome measure may overlook the practical difficulties in generating an evidence-base to demonstrate therapeutic efficacy and guide clinical management. The European Medicines Agency (EMA) recognises that the "conduct, analysis, and interpretation of studies in rare conditions…may be constrained to varying degrees by the prevalence of the disease" [160] and acknowledges the difficulty in demonstrating that a surrogate end-point adequately reflects the true clinical end-point.

The validation of CT densitometry as a surrogate measure of emphysema severity has followed a logical process as far as circumstances permit and represents a paradigm for other imaging biomarkers. The pathological definition of emphysema has meant, quite appropriately, that studies initially focused on comparison between densitometric indices and morphometry of resected pathological specimens [22, 26–28]; cross-sectional studies have compared CT indices with physiology (static lung function [161–163], exercise physiology [164]) and health status indices [163]); CT indices have since been shown to predict mortality in usual COPD [67]. In α_1-ATD, CT is a better predictor of mortality than physiology or health status [67, 165], a more sensitive measure of disease progression than health status and physiology [166], and CT progression correlates with deterioration in health status [167] and with FEV1 decline [168]. As a consequence of the published evidence, a meeting of the Blood Products Advisory Committee of the FDA [169] concluded that it accepted "serial lung density measurements by (high-resolution CT) as an appropriate clinically meaningful endpoint to assess the efficacy of augmentation therapy with (*i.v.* α_1-proteinase inhibitor) products on emphysema disease progression" and has permitted its use as a primary end-point in phase IV studies.

Interventional studies consistently demonstrate that a difference can be identified in the rate of decline in CT-derived lung density measurements in association with α_1-AT augmentation therapy [23, 31–33]. Despite the stance of the FDA [169] and EMA [160], the jury is still out over whether this alone is sufficient proof of clinical efficacy. There are some unresolved technical issues relating

to the finer points of CT densitometry but the crux of the matter comes down to whether CT densitometry can be viewed as a clinically meaningful outcome measure. The acceptance of CT as a valid biomarker was initially tempered by the requirement that structural changes would only be accepted as evidence of therapeutic efficacy if associated with functional improvement. Rather than recognising that a beneficial change in FEV_1 would not be seen in a study that was underpowered for demonstrating a change in FEV_1, unless by chance, the lack of signal seen in FEV_1 has been interpreted as indicating a lack of treatment effect. While the circular and illogical nature of this argument is self-evident, the solution to address this impasse is not. The acceptance or rejection of the findings of the RAPID study [33] as a clinically meaningful outcome may determine future strategy for the development of quantitative imaging, since the imaging methodology that was employed in this study is the most comprehensively validated to date. Current approaches to the validation of novel imaging biomarkers do not appear to be substantially different to that used in CT densitometry. Consequently, unless *a priori* strategic planning occurs between clinicians, scientists, industry, and regulatory and reimbursement bodies, the effective application of these new techniques for patient benefit may ultimately not be realised.

Conclusion

Imaging techniques have expanded from the realm of pure diagnostics to metrics in response to the need for patient-centred outcomes that are noninvasive, repeatable, sensitive and specific. However, consensus on appropriate validation and standardisation of these novel methodologies seems hard to achieve. Furthermore, the apparent reluctance of regulatory and reimbursement authorities to accept these new measures as clinically meaningful outcomes continues to pose a challenge to clinicians and the wider research community. In order for patients to receive prompt benefit from the considerable technical advances that have occurred, future research strategy should be led by physicians with an understanding of the clinical aspects of respiratory disease, in close communication with patients, industry and the regulators.

Statement of Interest

D.G. Parr reports personal fees from Grifols/Talecris Biotherapeutics and Hoffmann–La Roche for previous work as an independent consultant and his spouse has previously worked for Grifols/Talecris Biotherapeutics.

References

1. van Beek EJ, Hoffman EA. Functional imaging: CT and MRI. *Clin Chest Med* 2008; 29: 195–216.
2. Washko GR, Parraga G, Coxson HO. Quantitative pulmonary imaging using computed tomography and magnetic resonance imaging. *Respirology* 2012; 17: 432–444.
3. Hoffman EA, Chon D. Computed tomography studies of lung ventilation and perfusion. *Proc Am Thorac Soc* 2005; 2: 492–498.
4. Fain SB, Korosec FR, Holmes JH, *et al.* Functional lung imaging using hyperpolarized gas MRI. *J Magn Reson Imaging* 2007; 25: 910–923.
5. Washko GR. The role and potential of imaging in COPD. *Med Clin North Am* 2012; 96: 729–743.
6. Dolovich M, Labiris R. Imaging drug delivery and drug responses in the lung. *Proc Am Thorac Soc* 2004; 1: 329–337.
7. Hsia CC, Hyde DM, Ochs M, *et al.* An official research policy statement of the American Thoracic Society/European Respiratory Society: standards for quantitative assessment of lung structure. *Am J Respir Crit Care Med* 2010; 181: 394–418.
8. Thurlbeck WM. Measurement of pulmonary emphysema. *Am Rev Respir Dis* 1967; 95: 752–764.
9. Kalender WA, Rienmuller R, Seissler W, *et al.* Measurement of pulmonary parenchymal attenuation: use of spirometric gating with quantitative CT. *Radiology* 1990; 175: 265–268.
10. Moroni C, Mascalchi M, Camiciottoli G, *et al.* Comparison of spirometric-gated and -ungated HRCT in COPD. *J Comput Assist Tomogr* 2003; 27: 375–379.
11. Hoffman EA, Reinhardt JM, Sonka M, *et al.* Characterization of the interstitial lung diseases *via* density-based and texture-based analysis of computed tomography images of lung structure and function. *Acad Radiol* 2003; 10: 1104–1118.

12. Stolk J, Dirksen A, van der Lugt AA, *et al.* Repeatability of lung density measurements with low-dose computed tomography in subjects with alpha-1-antitrypsin deficiency-associated emphysema. *Invest Radiol* 2001; 36: 648–651.
13. Dirksen A. Monitoring the progress of emphysema by repeat computed tomography scans with focus on noise reduction. *Proc Am Thorac Soc 2008 15*, 5: 925–928.
14. Stoel BC, Putter H, Bakker ME, *et al.* Volume correction in computed tomography densitometry for follow-up studies on pulmonary emphysema. *Proc Am Thorac Soc* 2008; 5: 919–924.
15. Shaker SB, Dirksen A, Laursen LC, *et al.* Volume adjustment of lung density by computed tomography scans in patients with emphysema. *Acta Radiol* 2004; 45: 417–423.
16. Coxson HO. Quantitative chest tomography in COPD research: chairman's summary. *Proc Am Thorac Soc* 2008; 5: 874–877.
17. Washko GR, Hunninghake GM, Fernandez IE, *et al.* Lung volumes and emphysema in smokers with interstitial lung abnormalities. *N Engl J Med* 2011; 364: 897–906.
18. Wilson AG, Massarella GR, Pride NB. Elastic properties of airways in human lungs *post mortem*. *Am Rev Respir Dis* 1974; 110: 716–729.
19. Diaz AA, Han MK, Come CE, *et al.* Effect of emphysema on CT scan measures of airway dimensions in smokers. *Chest* 2013; 143: 687–693.
20. Scichilone N, Bruno A, Marchese R, *et al.* Association between reduced bronchodilatory effect of deep inspiration and loss of alveolar attachments. *Respir Res* 2005; 6: 55.
21. Bakker ME, Stolk J, Reiber JH, *et al.* Influence of inspiration level on bronchial lumen measurements with computed tomography. *Respir Med* 2012; 106: 677–686.
22. Hruban RH, Meziane MA, Zerhouni EA, *et al.* High resolution computed tomography of inflation-fixed lungs. Pathologic-radiologic correlation of centrilobular emphysema. *Am Rev Respir Dis* 1987; 136: 935–940.
23. Parr DG, Dirksen A, Piitulainen E, *et al.* Exploring the optimum approach to the use of CT densitometry in a randomised placebo-controlled study of augmentation therapy in alpha 1-antitrypsin deficiency. *Respir Res* 2009; 10: 75.
24. Hayhurst MD, MacNee W, Flenley DC, *et al.* Diagnosis of pulmonary emphysema by computerised tomography. *Lancet* 1984; 2: 320–322.
25. Gould GA, MacNee W, McLean A, *et al.* CT measurements of lung density in life can quantitate distal airspace enlargement – an essential defining feature of human emphysema. *Am Rev Respir Dis* 1988; 137: 380–392.
26. Gevenois PA, de Maertelaer V, De Vuyst P, *et al.* Comparison of computed density and macroscopic morphometry in pulmonary emphysema. *Am J Respir Crit Care Med* 1995; 152: 653–657.
27. Gevenois PA, De Vuyst P, de Maertelaer V, *et al.* Comparison of computed density and microscopic morphometry in pulmonary emphysema. *Am J Respir Crit Care Med* 1996; 154: 187–192.
28. Muller NL, Staples CA, Miller RR, *et al.* "Density mask". An objective method to quantitate emphysema using computed tomography. *Chest* 1988; 94: 782–787.
29. Yuan R, Mayo JR, Hogg JC, *et al.* The effects of radiation dose and CT manufacturer on measurements of lung densitometry. *Chest* 2007; 132: 617–623.
30. Schluchter MD, Stoller JK, Barker AF, *et al.* Feasibility of a clinical trial of augmentation therapy for α_1-antitrypsin deficiency. The Alpha 1-Antitrypsin Deficiency Registry Study Group. *Am J Respir Crit Care Med* 2000; 161: 796–801.
31. Dirksen A, Dijkman JH, Madsen F, *et al.* A randomized clinical trial of α_1-antitrypsin augmentation therapy. *Am J Respir Crit Care Med* 1999; 160: 1468–1472.
32. Dirksen A, Piitulainen E, Parr DG, *et al.* Exploring the role of CT densitometry: a randomised study of augmentation therapy in alpha1-antitrypsin deficiency. *Eur Respir J* 2009; 33: 1345–1353.
33. Chapman KR, Burdon JGW, Piitulainen E, *et al.* IV alpha1 antitrypsin (A1AT) preserves lung density in homozygous alpha1 antitrypsin deficiency (A1ATD): a randomized, placebo-controlled trial. *Am J Respir Crit Care Med* 2013; 187: Suppl., A45385.
34. Stolk J, Stockley RA, Stoel BC, *et al.* Randomised controlled trial for emphysema with a selective agonist of the γ-type retinoic acid receptor. *Eur Respir J* 2012; 40: 306–312.
35. Parr DG, Sevenoaks M, Deng C, *et al.* Detection of emphysema progression in alpha 1-antitrypsin deficiency using CT densitometry; methodological advances. *Respir Res* 2008; 9: 21.
36. Coxson HO, Nasute Fauerbach PV, Storness-Bliss C, *et al.* Computed tomography assessment of lung volume changes after bronchial valve treatment. *Eur Respir J* 2008; 32: 1443–1450.
37. Sterman DH, Mehta AC, Wood DE, *et al.* A multicenter pilot study of a bronchial valve for the treatment of severe emphysema. *Respiration* 2010; 79: 222–233.
38. Uppaluri R, Hoffman EA, Sonka M, *et al.* Computer recognition of regional lung disease patterns. *Am J Respir Crit Care Med* 1999; 160: 648–654.
39. Uppaluri R, Hoffman EA, Sonka M, *et al.* Interstitial lung disease: A quantitative study using the adaptive multiple feature method. *Am J Respir Crit Care Med* 1999; 159: 519–525.
40. Uppaluri R, Mitsa T, Sonka M, *et al.* Quantification of pulmonary emphysema from lung computed tomography images. *Am J Respir Crit Care Med* 1997; 156: 248–254.

41. Mishima M, Hirai T, Itoh H, *et al.* Complexity of terminal airspace geometry assessed by lung computed tomography in normal subjects and patients with chronic obstructive pulmonary disease. *Proc Natl Acad Sci USA* 1999; 96: 8829–8834.
42. Xu Y, Sonka M, McLennan G, *et al.* MDCT-based 3-D texture classification of emphysema and early smoking related lung pathologies. *IEEE Trans Med Imaging* 2006; 25: 464–475.
43. Sorensen L, Shaker SB, de Bruijne M. Texture classification in lung CT using local binary patterns. *Med Image Comput Comput Assist Interv* 2008; 11: 934–941.
44. Woodhouse N, Wild JM, Paley MN, *et al.* Combined helium-3/proton magnetic resonance imaging measurement of ventilated lung volumes in smokers compared to never-smokers. *J Magn Reson Imaging* 2005; 21: 365–369.
45. van Beek EJ, Wild JM, Kauczor HU, *et al.* Functional MRI of the lung using hyperpolarized 3-helium gas. *J Magn Reson Imaging* 2004; 20: 540–554.
46. Kirby M, Heydarian M, Svenningsen S, *et al.* Hyperpolarized ^{3}He magnetic resonance functional imaging semiautomated segmentation. *Acad Radiol* 2012; 19: 141–152.
47. Tustison NJ, Avants BB, Flors L, *et al.* Ventilation-based segmentation of the lungs using hyperpolarized ^{3}He MRI. *J Magn Reson Imaging* 2011; 34: 831–841.
48. Woods JC, Choong CK, Yablonskiy DA, *et al.* Hyperpolarized ^{3}He diffusion MRI and histology in pulmonary emphysema. *Magn Reson Med* 2006; 56: 1293–1300.
49. Jacob RE, Minard KR, Laicher G, *et al.* 3D ^{3}He diffusion MRI as a local *in vivo* morphometric tool to evaluate emphysematous rat lungs. *J Appl Physiol* 2008; 105: 1291–1300.
50. Diaz S, Casselbrant I, Piitulainen E, *et al.* Validity of apparent diffusion coefficient hyperpolarized ^{3}He-MRI using MSCT and pulmonary function tests as references. *Eur J Radiol* 2009; 71: 257–263.
51. Kirby M, Owrangi A, Svenningsen S, *et al.* On the role of abnormal $D_{L,CO}$ in ex-smokers without airflow limitation: symptoms, exercise capacity and hyperpolarised helium-3 MRI. *Thorax* 2013; 68: 752–759.
52. Mathew L, Evans A, Ouriadov A, *et al.* Hyperpolarized ^{3}He magnetic resonance imaging of chronic obstructive pulmonary disease: reproducibility at 3.0 tesla. *Acad Radiol* 2008; 15: 1298–1311.
53. Kirby M, Mathew L, Wheatley A, *et al.* Chronic obstructive pulmonary disease: longitudinal hyperpolarized ^{3}He MR imaging. *Radiology* 2010; 256: 280–289.
54. Woods JC, Yablonskiy DA, Choong CK, *et al.* Long-range diffusion of hyperpolarized 3He in explanted normal and emphysematous human lungs *via* magnetization tagging. *J Appl Physiol* 2005; 99: 1992–1997.
55. Wang C, Miller GW, Altes TA, *et al.* Time dependence of ^{3}He diffusion in the human lung: measurement in the long-time regime using stimulated echoes. *Magn Reson Med* 2006; 56: 296–309.
56. Kirby M, Heydarian M, Wheatley A, *et al.* Evaluating bronchodilator effects in chronic obstructive pulmonary disease using diffusion-weighted hyperpolarized helium-3 magnetic resonance imaging. *J Appl Physiol* 2012; 112: 651–657.
57. Ohno Y, Hatabu H. Basics concepts and clinical applications of oxygen-enhanced MR imaging. *Eur J Radiol* 2007; 64: 320–328.
58. Theilmann RJ, Arai TJ, Samiee A, *et al.* Quantitative MRI measurement of lung density must account for the change in T_2* with lung inflation. *J Magn Reson Imaging* 2009; 30: 527–534.
59. Hogg JC, Macklem PT, Thurlbeck WM. Site and nature of airway obstruction in chronic obstructive lung disease. *N Engl J Med* 1968; 278: 1355–1360.
60. Nakano Y, Wong JC, de Jong PA, *et al.* The prediction of small airway dimensions using computed tomography. *Am J Respir Crit Care Med* 2005; 171: 142–146.
61. de Jong PA, Muller NL, Pare PD, *et al.* Computed tomographic imaging of the airways: relationship to structure and function. *Eur Respir J* 2005; 26: 140–152.
62. Hasegawa M, Nasuhara Y, Onodera Y, *et al.* Airflow limitation and airway dimensions in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2006; 173: 1309–1315.
63. Nakano Y, Muro S, Sakai H, *et al.* Computed tomographic measurements of airway dimensions and emphysema in smokers. Correlation with lung function. *Am J Respir Crit Care Med* 2000; 162: 1102–1108.
64. Kim N, Seo JB, Song KS, *et al.* Semi-automatic measurement of the airway dimension by computed tomography using the full-with-half-maximum method: a study of the measurement accuracy according to the orientation of an artificial airway. *Korean J Radiol* 2008; 9: 236–242.
65. Estepar RS, Washko GG, Silverman EK, *et al.* Accurate airway wall estimation using phase congruency. *Med Image Comput Comput Assist Interv* 2006; 9: 125–134.
66. Diaz AA, Morales A, Diaz JC, *et al.* CT and physiologic determinants of dyspnea and exercise capacity during the six-minute walk test in mild COPD. *Respir Med* 2013; 107: 570–579.
67. Johannessen A, Skorge TD, Bottai M, *et al.* Mortality by level of emphysema and airway wall thickness. *Am J Respir Crit Care Med* 2013; 187: 602–608.
68. Hogg JC, McDonough JE, Sanchez PG, *et al.* Micro-computed tomography measurements of peripheral lung pathology in chronic obstructive pulmonary disease. *Proc Am Thorac Soc* 2009; 6: 546–549.
69. Diaz AA, Valim C, Yamashiro T, *et al.* Airway count and emphysema assessed by chest CT imaging predicts clinical outcome in smokers. *Chest* 2010; 138: 880–887.
70. Orlandi I, Moroni C, Camiciottoli G, *et al.* Chronic obstructive pulmonary disease: thin-section CT measurement of airway wall thickness and lung attenuation. *Radiology* 2005; 234: 604–610.

71. Patel BD, Coxson HO, Pillai SG, *et al.* Airway wall thickening and emphysema show independent familial aggregation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2008; 178: 500–505.
72. Dame Carroll JR, Chandra A, Jones AS, *et al.* Airway dimensions measured from micro-computed tomography and high-resolution computed tomography. *Eur Respir J* 2006; 28: 712–720.
73. Achenbach T, Weinheimer O, Brochhausen C, *et al.* Accuracy of automatic airway morphometry in computed tomography-correlation of radiological-pathological findings. *Eur J Radiol* 2012; 81: 183–188.
74. Montaudon M, Berger P, de Dietrich G, *et al.* Assessment of airways with three-dimensional quantitative thin-section CT: *in vitro* and *in vivo* validation. *Radiology* 2007; 242: 563–572.
75. Montaudon M, Desbarats P, Berger P, *et al.* Assessment of bronchial wall thickness and lumen diameter in human adults using multi-detector computed tomography: comparison with theoretical models. *J Anat* 2007; 211: 579–588.
76. Yamashiro T, Matsuoka S, Estepar RS, *et al.* Quantitative assessment of bronchial wall attenuation with thin-section CT: An indicator of airflow limitation in chronic obstructive pulmonary disease. *AJR Am J Roentgenol* 2010; 195: 363–369.
77. Washko GR, Dransfield MT, Estepar RS, *et al.* Airway wall attenuation: a biomarker of airway disease in subjects with COPD. *J Appl Physiol* 2009; 107: 185–191.
78. Busacker A, Newell JD Jr, Keefe T, *et al.* A multivariate analysis of risk factors for the air-trapping asthmatic phenotype as measured by quantitative CT analysis. *Chest* 2009; 135: 48–56.
79. Hersh CP, Washko GR, Estepar RS, *et al.* Paired inspiratory-expiratory chest CT scans to assess for small airways disease in COPD. *Respir Res* 2013; 14: 42.
80. Wiggs BR, Bosken C, Pare PD, *et al.* A model of airway narrowing in asthma and in chronic obstructive pulmonary disease. *Am Rev Respir Dis* 1992; 145: 1251–1258.
81. Niimi A, Matsumoto H, Amitani R, *et al.* Airway wall thickness in asthma assessed by computed tomography. Relation to clinical indices. *Am J Respir Crit Care Med* 2000; 162: 1518–1523.
82. Little SA, Sproule MW, Cowan MD, *et al.* High resolution computed tomographic assessment of airway wall thickness in chronic asthma: reproducibility and relationship with lung function and severity. *Thorax* 2002; 57: 247–253.
83. Matsumoto H, Niimi A, Tabuena RP, *et al.* Airway wall thickening in patients with cough variant asthma and nonasthmatic chronic cough. *Chest* 2007; 131: 1042–1049.
84. Hoshino M, Ohtawa J. Effects of adding omalizumab, an anti-immunoglobulin E antibody, on airway wall thickening in asthma. *Respiration* 2012; 83: 520–528.
85. Hoshino M, Ohtawa J. Effects of budesonide/formoterol combination therapy *versus* budesonide alone on airway dimensions in asthma. *Respirology* 2012; 17: 639–646.
86. Hoshino M, Matsuoka S, Handa H, *et al.* Correlation between airflow limitation and airway dimensions assessed by multidetector CT in asthma. *Respir Med* 2010; 104: 794–800.
87. Zeidler MR, Kleerup EC, Goldin JG, *et al.* Montelukast improves regional air-trapping due to small airways obstruction in asthma. *Eur Respir J* 2006; 27: 307–315.
88. Goldin JG, Tashkin DP, Kleerup EC, *et al.* Comparative effects of hydrofluoroalkane and chlorofluorocarbon beclomethasone dipropionate inhalation on small airways: assessment with functional helical thin-section computed tomography. *J Allergy Clin Immunol* 1999; 104: S258–267.
89. de Lange EE, Altes TA, Patrie JT, *et al.* Evaluation of asthma with hyperpolarized helium-3 MRI: correlation with clinical severity and spirometry. *Chest* 2006; 130: 1055–1062.
90. Samee S, Altes T, Powers P, *et al.* Imaging the lungs in asthmatic patients by using hyperpolarized helium-3 magnetic resonance: assessment of response to methacholine and exercise challenge. *J Allergy Clin Immunol* 2003; 111: 1205–1211.
91. de Lange EE, Altes TA, Patrie JT, *et al.* Changes in regional airflow obstruction over time in the lungs of patients with asthma: evaluation with ^{3}He MR imaging. *Radiology* 2009; 250: 567–575.
92. Svenningsen S, Kirby M, Starr D, *et al.* Hyperpolarized He and Xe MRI: Differences in asthma before bronchodilation. *J Magn Reson Imaging* 2013 [In press DOI: 10.1002/jmri.24111].
93. Loeve M, van Hal PT, Robinson P, *et al.* The spectrum of structural abnormalities on CT scans from patients with CF with severe advanced lung disease. *Thorax* 2009; 64: 876–882.
94. Long FR, Williams RS, Castile RG. Structural airway abnormalities in infants and young children with cystic fibrosis. *J Pediatr* 2004; 144: 154–161.
95. Brody AS. Early morphologic changes in the lungs of asymptomatic infants and young children with cystic fibrosis. *J Pediatr* 2004; 144: 145–146.
96. Judge EP, Dodd JD, Masterson JB, *et al.* Pulmonary abnormalities on high-resolution CT demonstrate more rapid decline than FEV_1 in adults with cystic fibrosis. *Chest* 2006; 130: 1424–1432.
97. de Jong PA, Nakano Y, Lequin MH, *et al.* Progressive damage on high resolution computed tomography despite stable lung function in cystic fibrosis. *Eur Respir J* 2004; 23: 93–97.
98. Martinez TM, Llapur CJ, Williams TH, *et al.* High-resolution computed tomography imaging of airway disease in infants with cystic fibrosis. *Am J Respir Crit Care Med* 2005; 172: 1133–1138.
99. Robinson TE, Goris ML, Zhu HJ, *et al.* Dornase alfa reduces air trapping in children with mild cystic fibrosis lung disease: a quantitative analysis. *Chest* 2005; 128: 2327–2335.

100. Brody AS, Klein JS, Molina PL, *et al.* High-resolution computed tomography in young patients with cystic fibrosis: distribution of abnormalities and correlation with pulmonary function tests. *J Pediatr* 2004; 145: 32–38.
101. Bhalla M, Turcios N, Aponte V, *et al.* Cystic fibrosis: scoring system with thin-section CT. *Radiology* 1991; 179: 783–788.
102. Nasr SZ, Kuhns LR, Brown RW, *et al.* Use of computerized tomography and chest x-rays in evaluating efficacy of aerosolized recombinant human DNase in cystic fibrosis patients younger than age 5 years: a preliminary study. *Pediatr Pulmonol* 2001; 31: 377–382.
103. de Jong PA, Ottink MD, Robben SG, *et al.* Pulmonary disease assessment in cystic fibrosis: comparison of CT scoring systems and value of bronchial and arterial dimension measurements. *Radiology* 2004; 231: 434–439.
104. Brody AS, Kosorok MR, Li Z, *et al.* Reproducibility of a scoring system for computed tomography scanning in cystic fibrosis. *J Thorac Imaging* 2006; 21: 14–21.
105. de Jong PA, Lindblad A, Rubin L, *et al.* Progression of lung disease on computed tomography and pulmonary function tests in children and adults with cystic fibrosis. *Thorax* 2006; 61: 80–85.
106. Quan JM, Tiddens HA, Sy JP, *et al.* A two-year randomized, placebo-controlled trial of dornase alfa in young patients with cystic fibrosis with mild lung function abnormalities. *J Pediatr* 2001; 139: 813–820.
107. Robinson TE, Leung AN, Northway WH, *et al.* Composite spirometric-computed tomography outcome measure in early cystic fibrosis lung disease. *Am J Respir Crit Care Med* 2003; 168: 588–593.
108. Robinson TE, Leung AN, Northway WH, *et al.* Spirometer-triggered high-resolution computed tomography and pulmonary function measurements during an acute exacerbation in patients with cystic fibrosis. *J Pediatr* 2001; 138: 553–559.
109. Shah RM, Sexauer W, Ostrum BJ, *et al.* High-resolution CT in the acute exacerbation of cystic fibrosis: evaluation of acute findings, reversibility of those findings, and clinical correlation. *Am J Roentgenol* 1997; 169: 375–380.
110. Brody AS, Molina PL, Klein JS, *et al.* High-resolution computed tomography of the chest in children with cystic fibrosis: support for use as an outcome surrogate. *Pediatr Radiol* 1999; 29: 731–735.
111. Tiddens HA, de Jong PA. Imaging and clinical trials in cystic fibrosis. *Proc Am Thorac Soc* 2007; 4: 343–346.
112. Kang EY, Miller RR, Muller NL. Bronchiectasis: comparison of preoperative thin-section CT and pathologic findings in resected specimens. *Radiology* 1995; 195: 649–654.
113. Davis SD, Fordham LA, Brody AS, *et al.* Computed tomography reflects lower airway inflammation and tracks changes in early cystic fibrosis. *Am J Respir Crit Care Med* 2007; 175: 943–950.
114. Brody AS, Sucharew H, Campbell JD, *et al.* Computed tomography correlates with pulmonary exacerbations in children with cystic fibrosis. *Am J Respir Crit Care Med* 2005; 172: 1128–1132.
115. Loeve M, Hop WC, de Bruijne M, *et al.* Chest computed tomography scores are predictive of survival in patients with cystic fibrosis awaiting lung transplantation. *Am J Respir Crit Care Med* 2012; 185: 1096–1103.
116. Long FR. High-resolution computed tomography of the lung in children with cystic fibrosis: technical factors. *Proc Am Thorac Soc* 2007; 4: 306–309.
117. de Jong PA, Owens CM. Radiation dose for pediatric patients with cystic fibrosis: a continuous adjustment process and remaining concern. *Chest* 2012; 142: 1077.
118. Mathews JD, Forsythe AV, Brady Z, *et al.* Cancer risk in 680 000 people exposed to computed tomography scans in childhood or adolescence: data linkage study of 11 million Australians. *BMJ* 2013; 346: f2360.
119. Failo R, Wielopolski PA, Tiddens HA, *et al.* Lung morphology assessment using MRI: a robust ultra-short TR/TE 2D steady state free precession sequence used in cystic fibrosis patients. *Magn Reson Med* 2009; 61: 299–306.
120. Puderbach M, Eichinger M, Gahr J, *et al.* Proton MRI appearance of cystic fibrosis: comparison to CT. *Eur Radiol* 2007; 17: 716–724.
121. Puderbach M, Eichinger M, Haeselbarth J, *et al.* Assessment of morphological MRI for pulmonary changes in cystic fibrosis (CF) patients: comparison to thin-section CT and chest x-ray. *Invest Radiol* 2007; 42: 715–725.
122. McMahon CJ, Dodd JD, Hill C, *et al.* Hyperpolarized 3helium magnetic resonance ventilation imaging of the lung in cystic fibrosis: comparison with high resolution CT and spirometry. *Eur Radiol* 2006; 16: 2483–2490.
123. Woodhouse N, Wild JM, van Beek EJ, *et al.* Assessment of hyperpolarized ^{3}He lung MRI for regional evaluation of interventional therapy: a pilot study in pediatric cystic fibrosis. *J Magn Reson Imaging* 2009; 30: 981–988.
124. Eichinger M, Puderbach M, Fink C, *et al.* Contrast-enhanced 3D MRI of lung perfusion in children with cystic fibrosis – initial results. *Eur Radiol* 2006; 16: 2147–2152.
125. Jakob PM, Wang T, Schultz G, *et al.* Assessment of human pulmonary function using oxygen-enhanced T_1 imaging in patients with cystic fibrosis. *Magn Reson Med* 2004; 51: 1009–1016.
126. Mentore K, Froh DK, de Lange EE, *et al.* Hyperpolarized HHe 3 MRI of the lung in cystic fibrosis: assessment at baseline and after bronchodilator and airway clearance treatment. *Acad Radiol* 2005; 12: 1423–1429.
127. van Beek EJ, Hill C, Woodhouse N, *et al.* Assessment of lung disease in children with cystic fibrosis using hyperpolarized 3-helium MRI: comparison with Shwachman score, Chrispin-Norman score and spirometry. *Eur Radiol* 2007; 17: 1018–1024.
128. Ohno Y, Koyama H, Nogami M, *et al.* Dynamic perfusion MRI: capability for evaluation of disease severity and progression of pulmonary arterial hypertension in patients with connective tissue disease. *J Magn Reson Imaging* 2008; 28: 887–899.

129. Park SO, Seo JB, Kim N, *et al.* Comparison of usual interstitial pneumonia and nonspecific interstitial pneumonia: quantification of disease severity and discrimination between two diseases on HRCT using a texture-based automated system. *Korean J Radiol* 2011; 12: 297–307.
130. Yoon RG, Seo JB, Kim N, *et al.* Quantitative assessment of change in regional disease patterns on serial HRCT of fibrotic interstitial pneumonia with texture-based automated quantification system. *Eur Radiol* 2013; 23: 692–701.
131. Erdal BS, Crouser ED, Yildiz V, *et al.* Quantitative computerized two-point correlation analysis of lung CT scans correlates with pulmonary function in pulmonary sarcoidosis. *Chest* 2012; 142: 1589–1597.
132. Yao J, Taveira-DaSilva AM, Colby TV, *et al.* CT grading of lung disease in lymphangioleiomyomatosis. *Am J Roentgenol* 2012; 199: 787–793.
133. Jones HA, Sriskandan S, Peters AM, *et al.* Dissociation of neutrophil emigration and metabolic activity in lobar pneumonia and bronchiectasis. *Eur Respir J* 1997; 10: 795–803.
134. Subramanian DR, Jenkins L, Edgar R, *et al.* Assessment of pulmonary neutrophilic inflammation in emphysema by quantitative positron emission tomography. *Am J Respir Crit Care Med* 2012; 186: 1125–1132.
135. Jones HA, Clark RJ, Rhodes CG, *et al. In vivo* measurement of neutrophil activity in experimental lung inflammation. *Am J Respir Crit Care Med* 1994; 149: 1635–1639.
136. Jones HA, Boobis AR, Hamacher K, *et al.* PET imaging of pulmonary fibrosis. *J Nucl Med* 2003; 44: 483–484.
137. Jones HA, Marino PS, Shakur BH, *et al. In vivo* assessment of lung inflammatory cell activity in patients with COPD and asthma. *Eur Respir J* 2003; 21: 567–573.
138. Currie DC, Saverymuttu SH, Peters AM, *et al.* Indium-111-labelled granulocyte accumulation in respiratory tract of patients with bronchiectasis. *Lancet* 1987; 1: 1335–1339.
139. Amin R, Charron M, Grinblat L, *et al.* Cystic fibrosis: detecting changes in airway inflammation with FDG PET/CT. *Radiology* 2012; 264: 868–875.
140. Klein M, Cohen-Cymberknoh M, Armoni S, *et al.* ^{18}F-fluorodeoxyglucose-PET/CT imaging of lungs in patients with cystic fibrosis. *Chest* 2009; 136: 1220–1228.
141. Taylor IK, Hill AA, Hayes M, *et al.* Imaging allergen-invoked airway inflammation in atopic asthma with [^{18}F]-fluorodeoxyglucose and positron emission tomography. *Lancet* 1996; 347: 937–940.
142. Lewis PJ, Salama A. Uptake of fluorine-18-fluorodeoxyglucose in sarcoidosis. *J Nucl Med* 1994; 35: 1647–1649.
143. Brudin LH, Valind SO, Rhodes CG, *et al.* Fluorine-18 deoxyglucose uptake in sarcoidosis measured with positron emission tomography. *Eur J Nucl Med* 1994; 21: 297–305.
144. Keijsers RG, Verzijlbergen EJ, van den Bosch JM, *et al.* ^{18}F-FDG PET as a predictor of pulmonary function in sarcoidosis. *Sarcoidosis Vasc Diffuse Lung Dis* 2011; 28: 123–129.
145. Groves AM, Win T, Screaton NJ, *et al.* Idiopathic pulmonary fibrosis and diffuse parenchymal lung disease: implications from initial experience with ^{18}F-FDG PET/CT. *J Nucl Med* 2009; 50: 538–545.
146. Win T, Screaton NJ, Porter J, *et al.* Novel positron emission tomography/computed tomography of diffuse parenchymal lung disease combining a labelled somatostatin receptor analogue and 2-deoxy-2[^{18}F]fluoro-D-glucose. *Mol Imaging* 2012; 11: 91–98.
147. Choi S, Hoffman EA, Wenzel SE, *et al.* Registration-based assessment of regional lung function *via* volumetric CT images of normals *vs.* severe asthmatics. *J Appl Physiol.* 2013; 115: 730–742.
148. Vidal Melo MF, Winkler T, Harris RS, *et al.* Spatial heterogeneity of lung perfusion assessed with ^{13}N PET as a vascular biomarker in chronic obstructive pulmonary disease. *J Nucl Med* 2010; 51: 57–65.
149. Vidal Melo MF, Layfield D, Harris RS, *et al.* Quantification of regional ventilation-perfusion ratios with PET. *J Nucl Med* 2003; 44: 1982–1991.
150. Yin Y, Choi J, Hoffman EA, *et al.* A multiscale MDCT image-based breathing lung model with time-varying regional ventilation. *J Comput Phys* 2013; 244: 168–192.
151. Kumar H, Tawhai MH, Hoffman EA, *et al.* The effects of geometry on airflow in the acinar region of the human lung. *J Biomech* 2009; 42: 1635–1642.
152. Lambert AR, O'Shaughnessy P, Tawhai MH, *et al.* Regional deposition of particles in an image-based airway model: large-eddy simulation and left-right lung ventilation asymmetry. *Aerosol Sci Technol* 2011; 45: 11–25.
153. De Backer LA, Vos WG, Salgado R, *et al.* Functional imaging using computer methods to compare the effect of salbutamol and ipratropium bromide in patient-specific airway models of COPD. *Int J Chron Obstruct Pulmon Dis* 2011; 6: 637–646.
154. De Backer JW, Vos WG, Vinchurkar SC, *et al.* Validation of computational fluid dynamics in CT-based airway models with SPECT/CT. *Radiology* 2010; 257: 854–862.
155. Kurland BF, Gerstner ER, Mountz JM, *et al.* Promise and pitfalls of quantitative imaging in oncology clinical trials. *Magn Reson Imaging* 2012; 30: 1301–1312.
156. Serkova NJ, Garg K, Bradshaw-Pierce EL. Oncologic imaging end-points for the assessment of therapy response. *Recent Pat Anticancer Drug Discov* 2009; 4: 36–53.
157. Miller MR, Hankinson J, Brusasco V, *et al.* Standardisation of spirometry. *Eur Respir J* 2005; 26: 319–338.
158. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. *Stat Med* 1989; 8: 431–440.
159. Food and Drug Administration Modernization Act of 1997, 21 USC § 301 (1997).
160. European Medicines Agency. Guidelines on Clinical Trials in Small Populations. www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003615.pdf Date last accessed: June 1, 2013. Date last updated: July 27, 2006.

161. Gould GA, Redpath AT, Ryan M, *et al.* Lung CT density correlates with measurements of airflow limitation and the diffusing capacity. *Eur Respir J* 1991; 4: 141–146.
162. Kinsella M, Muller NL, Abboud RT, *et al.* Quantitation of emphysema by computed tomography using a "density mask" program and correlation with pulmonary function tests. *Chest* 1990; 97: 315–321.
163. Dowson LJ, Guest PJ, Hill SL, *et al.* High-resolution computed tomography scanning in α1-antitrypsin deficiency: relationship to lung function and health status. *Eur Respir J* 2001; 17: 1097–1104.
164. Dowson LJ, Newall C, Guest PJ, *et al.* Exercise capacity predicts health status in α_1-antitrypsin deficiency. *Am J Respir Crit Care Med* 2001; 163: 936–941.
165. Dawkins PA, Dowson LJ, Guest PJ, *et al.* Predictors of mortality in α1-antitrypsin deficiency. *Thorax* 2003; 58: 1020–1026.
166. Dowson LJ, Guest PJ, Stockley RA. Longitudinal changes in physiological, radiological, and health status measurements in α_1-antitrypsin deficiency and factors associated with decline. *Am J Respir Crit Care Med* 2001; 164: 1805–1809.
167. Stolk J, Ng WH, Bakker ME, *et al.* Correlation between annual change in health status and computer tomography derived lung density in subjects with α1-antitrypsin deficiency. *Thorax* 2003; 58: 1027–1030.
168. Parr DG, Stoel BC, Stolk J, *et al.* Validation of computed tomographic lung densitometry for monitoring emphysema in α1-antitrypsin deficiency. *Thorax* 2006; 61: 485–490.
169. Blood Products Advisory Committee. 95th Meeting, July 20–21, 2009. www.fda.gov/downloads/advisorycommittees/committeesmeetingmaterials/bloodvaccinesandotherbiologics/bloodproductsadvisorycommittee/ucm171091.pdf Date last accessed: June 1, 2013. Date last updated: July 2009.

Chapter 13

Exercise tests

Thierry Troosters

SUMMARY: Exercise tests are increasingly used to assess the effects of pharmacological and nonpharmacological interventions. They are also used to study symptoms and the reduction thereof, as well as the mechanisms of action of interventions.

Laboratory tests are particularly useful to study the physiology of exercise, operational lung volumes, gas exchange and symptoms. Field tests are less complex and less controlled, but can be deployed in settings where an exercise laboratory is not available. Functional exercise tests are designed to assess the capacity of patients to cope with exercise relevant to daily function, such as walking, stair climbing or combined tasks. Cycle tests may be more sensitive to change in studies where besides lung function also peripheral muscle improves. Walking tests may be more sensitive to change when an intervention improves lung function only. A last choice in terms of determining the optimal test for a clinical trial is to use either endurance or incremental tests. In chronic obstructive pulmonary disease, endurance tests have shown to be much more responsive than incremental tests, particularly in patients who are ventilatory limited. Other factors that influence the response of interventions are the exercise intensity and the duration for which the intervention is applied.

Dept of Rehabilitation Sciences, Katholieke Universiteit Leuven, Leuven, and University Hospitals Leuven, Respiratory Division, Pulmonary Rehabilitation, Leuven, Belgium.

Correspondence: T. Troosters, University Hospitals Leuven, Respiratory Division, Pulmonary Rehabilitation, Herestraat 49, B3000 Leuven, Belgium.
Email: Thierry.troosters@med.kuleuven.be

Eur Respir Monogr 2013; 62: 170–178.
Copyright ERS 2013.
DOI: 10.1183/1025448x.10045612
Print ISBN: 978-1-84984-044-6
Online ISBN: 978-1-84984-045-3
Print ISSN: 1025-448x
Online ISSN: 2075-6674

Efficacy of interventions is assessed in clinical trials by evaluation of a single organ function (*e.g.* lung function or muscle strength), through integrated measures evaluating several organ functions at the same time (*e.g.* exercise testing) or through patient-reported outcomes (*e.g.* quality of life or symptoms). Exercise tests are a comprehensive way to assess patients' performance as it solicits several organ systems, including the lungs, the cardiovascular system and the peripheral muscle. Improving exercise tolerance is clinically more meaningful than improving lung function alone. Symptoms during exercise are one of the predominant features of chronic obstructive pulmonary disease (COPD) and can easily be assessed in a standardised way during exercise. Other typical outcomes of exercise assessment are the exercise performance, as assessed by endurance time or distance walked, and peak exercise tolerance. Finally, exercise tests can provide some information on the mechanism of action of interventions when, for example, cardiac function or ventilatory pattern are assessed during exercise. Although improved exercise tolerance is an important outcome, it can be assessed using several tests that address distinct exercise related features and, therefore, cannot be compared. For the purpose of this chapter, the commonly used exercise tests are classified in three (somewhat overlapping) types: 1) laboratory-based tests, field tests and functional tests; 2) walk tests and cycling tests; and 3) incremental tests and endurance tests.

In its draft guidance to the industry, the US Food and Drug Administration opened the door to exercise assessment showing efficacy as a secondary end-point [1]. From the text, however, it is clear that there is some reluctance to accept tests that are also dependent on motivation. This is obviously inherent to measures of exercise capacity. In June 2012, the European Medicines Agency stated that symptom relief during exercise or enhanced exercise capacity can be a co-primary end-point with lung function improvements [2]. Exercise assessment has found its way into contemporary pharmacological studies (typically in phase III or phase IV) but has been used for decades as a gold standard outcome to assess the effectiveness of pulmonary rehabilitation and surgical interventions, such as lung transplantation and volume reduction surgery.

Classification

Laboratory-based tests, field tests and functional tests

Laboratory based tests are typically conducted on a cycle ergometer or treadmill. They can apply incremental or constant work rates. Whereas the former seek to investigate the peak exercise performance and mechanisms limiting exercise, the latter apply a submaximal work rate and investigate how long this can be sustained. When constant work rate tests are conducted at relatively high work rates, the exercise-induced responses, in terms of oxygen consumption($V'O_2$), heart rate and particularly ventilation, may approach peak exercise responses but the approach to these peak responses is slower than in maximal exercise tests (fig. 1). Laboratory-based tests allow for the assessment of gas exchange, heart rate, and particularly operational lung volumes and symptoms at submaximal and maximal exercise tolerance. Laboratory tests allow the investigation of the mechanisms of action of interventions.

Field exercise tests are less sophisticated in terms of their setup and the equipment needed. Gas exchange measures can be performed using portable metabolic equipment but these tests are typically not meant for these assessments. Like laboratory-based tests, these tests can be incremental (incremental shuttle walk test [3]) or endurance tests (endurance shuttle walk tests [4]). The 6-min walk test (6MWT) [5] is a field test that has been popular for use in patients with respiratory diseases. Whereas in the shuttle walk tests, the time (or distance) is the outcome and speed is controlled, the time is controlled in timed walk tests such as the 6MWT.

Functional exercise tests are field tests that are more closely related to patient function: 6MWTs, stair tests, and combined tests such as the Glittre ADL test [6] or the Physical Functional

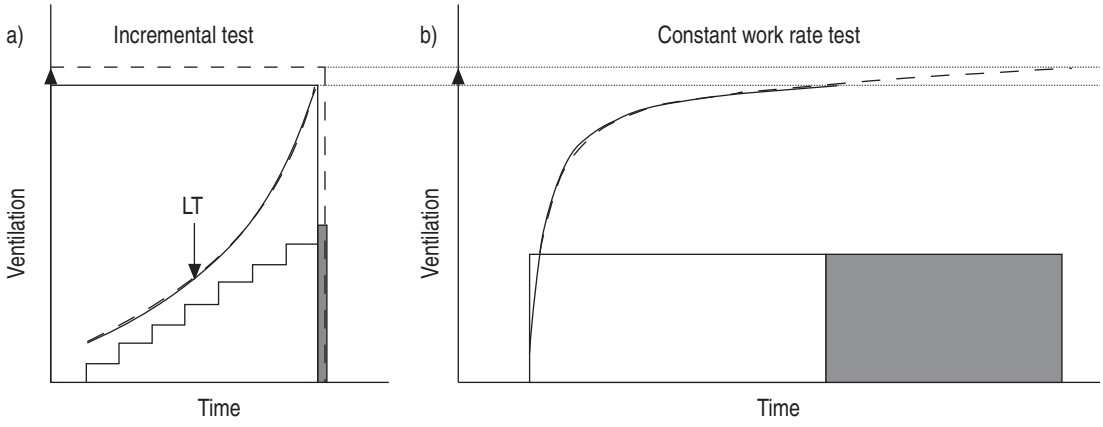


Figure 1. The theoretical effect of bronchodilator therapy (dashed line) on a) incremental and b) constant work rate tests in patients who are ventilatory limited. Bronchodilators increase peak ventilatory capacity. This results in a small effect in peak exercise tolerance, and a potentially much larger effect in endurance exercise capacity at high work rate (*e.g.* 75% of peak). The shaded area represents the expected improvement with an increased peak ventilation. LT: lactic threshold.

Performance 10 test [7]. Finally, a number of tests have recently been adopted from the geriatric field. These tests assess functional muscle strength and balance. An example is the sit-to-stand test. When the mechanism of action of a pharmacological intervention is to enhance exercise endurance, by improving the ventilatory capacity and hence oxygen delivery to the working muscle, it is unlikely to see functional tests that specifically solicit strength improvement. Such tests may, however, improve when interventions specifically target skeletal muscle function.

Walk *versus* cycling tests

One important paradigm to consider when selecting an exercise test is whether the test will use walking or cycling as a mode of exercise. Both forms of exercise have long been used as exercise modalities to elicit significant metabolic load on the exercising body. In the past few years, however, it has become clear that there are important differences between the two, which may make them more or less responsive to specific interventions. First, walking is a common exercise but it requires a lot of muscles to work against gravity. Work rate is difficult to standardise and is determined by body weight and incline. When walking on flat surface, work is very difficult to calculate and the effort of walking is determined by body weight, the type of surface (usually a corridor or treadmill) and the footwear of the subject [8]. Peak metabolic responses during walking are higher but the work is divided over more muscle groups, rendering walking exercises generally less fatiguing to the peripheral muscle [9]. Walking is associated with greater oxyhaemoglobin desaturation [10, 11] and with a later (or often even absent) onset of blood lactate accumulation compared with cycling [11]. Whereas during walking, dyspnoea is the predominant symptom at the end of exercise, cycling is characterised by a more variable symptom pattern including dyspnoea but also leg fatigue [12, 13]. Patients limited by breathlessness experience less quadriceps contractile fatigue with endurance exercise and seem to display more effect of bronchodilators [14]. As during walking, symptoms are more predominantly dyspnoea, it is no surprise that endurance walk tests showed to be more responsive than cycling to capture the effects of bronchodilators [15]. Tests of endurance, *i.e.* assessing the time until exhaustion, are suggested to be more responsive compared with tests with fixed walking times, such as the 6MWT [16]. Although theoretically, cycling tests may be more responsive to improvements after interventions that improve muscle function (*e.g.* exercise training), walking is also a responsive exercise modality in this setting, as walking exercises are typically included as training modality. Patients then have the additional benefits of enhanced mechanical walking efficiency, leading to enhanced walking performance.

Incremental *versus* endurance tests

Incremental exercise tests are designed to assess the patient's peak exercise capacity. Work rate is gradually increased until the patient cannot continue with the exercise. Work rate increments can occur every minute or continuously. The intension is to reach peak exercise tolerance within 7 to 12 min. Peak metabolic responses ($V'O_2$, ventilation and heart rate), more than peak work rate, are the typical outcomes during laboratory tests. The peak work rate obtained depends to a large extent on the work rate increment and, therefore, is a less optimal outcome [17]. The field test variant is the incremental shuttle walk test, where the time or distance covered are the classical outcomes. Peak exercise tests are as reliable in patients with COPD as in healthy subjects [18]. Similarly, in patients with pulmonary hypertension, peak $V'O_2$ showed a coefficient of variation of 5–6% [19]. The incremental shuttle walk test showed a systematic learning effect [20]. Incremental exercise tests bring patients to their physiological exercise limitation in a short time span. Pulmonary ventilation is accelerated when lactate is produced. Due to the rapid rise in pulmonary ventilation, the incremental exercise test is not a very responsive test in patients with ventilatory limitation (fig. 1a). One study that compared the responsiveness of several exercise tests to bronchodilator therapy showed that peak $V'O_2$ did not change significantly after treatment with oxitropium, whereas in these patients endurance time was enhanced by 19% [21]. Similarly, a

systematic review of the literature on pulmonary rehabilitation showed an increase in peak $V'O_2$ of 11% whereas endurance time after pulmonary rehabilitation increased by 87% [22].

Endurance tests typically impose a fixed work rate on patients after a resting period. Due to the single work rate condition, patients obtain steady state metabolic turnover, meaning that heart rate, $V'O_2$, carbon dioxide production ($V'CO_2$) and pulmonary ventilation reach steady-state values when the work rate is low. At high work rates, a continuous slow rise in metabolic parameters, particularly in $V'CO_2$ and pulmonary ventilation, is noted due to the lactate production. The speed with which the (near) steady-state conditions are reached depends on the physical fitness and the oxidative capacity of patients. This $V'O_2$ kinetic can be calculated as the mean response time, which is proportional to the time to reach the steady state (see [23] for a description of the calculation). In patients with a ventilatory limitation, the endurance test has the advantage that the ventilatory ceiling is approached more gently, leaving more time to continue exercising even if the improvement in peak ventilation is relatively small (fig. 1b). This makes the test more responsive to treatment. It can also be argued that submaximal, high-intensity endurance is a relevant exercise to test patients with COPD, as daily life in these patients if often conducted at relatively high work rates [24].

Commonly used tests in clinical trials

In the context of clinical trials, certain features of an exercise test are of utmost importance. Besides being a valid assessment of exercise tolerance, tests need to be reliable, reproducible in a multicentre setting and responsive to change. Ideally, the observed change should be meaningful to patients or in terms of potential morbidity benefits. Since exercise tests are complex routines, including patient and equipment preparation, equipment calibration, selection of the test protocol, and selection of the reported data, standardisation is of utmost importance in multicentre clinical trials. Exercise tests are widely applied in clinical routine but in the context of a multicentre clinical trial, the usual routine will have to be modified. This introduces a challenge to those designing and supervising trials. In many contemporary trials, formal training to ensure that the procedures are meticulously followed by study investigators and quality control is part of the good clinical practice procedures. In the next section, the aspects of standardisation that merit the attention of trial designers, as well as the responsiveness of the different tests, will be briefly highlighted.

Incremental laboratory-based exercise test

Incremental exercise tests are considered the gold standard for assessing peak exercise performance, and have diagnostic and predictive properties. Patients achieve near-maximal values for $V'O_2$. Gas exchange can be assessed at rest, during unloaded cycling and with increasing work rate until exhaustion. Indications and interpretation standards for the test are described in detail by a working group of the European Respiratory Society (ERS) [25] and a consensus document from the American Thoracic Society (ATS) and the American College of Chest Physicians [26]. Incremental exercise tests are often used in clinical practice for diagnostic purposes in clinical exercise laboratories. Diagnostic algorithms are published elsewhere [26] and software is available to help with the interpretation [27]. In COPD patients, these tests are not the most responsive in patients who are ventilatory limited, as ventilation increases rapidly towards the end of an incremental exercise test, leaving little room to improve unless major increases in the ventilatory capacity occur. After pulmonary rehabilitation, an improvement in peak work rate of ~10 W can be expected. With bronchodilators, the improvement is 3 to 7 W [28]. Although a minimally clinically important difference was suggested to be 4 W [29], this needs to be taken with a pinch of salt until these data are reproduced by others. 4 W is a very small difference, which may, at a group level, be indicative of overall response but should probably not be used in responder analysis, as a 4-W difference is hardly measurable and most reported incremental exercise protocols use 10-W·min^{-1} increments.

Nevertheless, an incremental exercise test is the underlying basis to conducting an endurance exercise test (see later). Clinical trials that use endurance exercise test results as an outcome measure should also include an incremental exercise test, at least during the baseline assessment.

Standardisation

Calibration of the gas and flow analysers is crucial but ensuring that the external load is correct is of crucial importance. Biological calibration, where the same subjects conduct a submaximal exercise on a regular basis, can be of help to detect calibration problems in real-life testing settings. Variation in $V'O_2$, $V'CO_2$ and heart rate should be less than 5% [30]. Standardising the laboratory environment in terms of temperature and humidity contribute to enhanced quality of the results obtained. An important aspect to control is the increase in work rate. Steeper increases in work rate lead to larger achieved peak work rates in COPD [17, 31] and pulmonary hypertension [32].

Endurance cycling test

Endurance tests are more responsive to interventions compared to incremental exercise tests. This is particularly so in patients who are ventilatory limited during exercise, as during an endurance test, the ventilation only slowly approaches the peak ventilation, leaving more opportunity to change. When patients are limited by peripheral muscle fatigue, the endurance test will not be sensitive to change after bronchodilator administration [33]. The time to exhaustion is relatively reproducible. A learning effect of 6% can be expected in a multicentre setting [34], so a practice endurance test may minimise the noise of the test. Endurance tests also allow assessment of operating lung volumes during exercise by performing repeated inspiratory capacity manoeuvres or using optoelectronic plethysmography (not feasible in multicentre studies). Inspiratory capacity manoeuvres can be performed reliably and in multicentre studies [34]. Symptom perception can also be assessed during the endurance test, offering unique integration of exercise, respiratory mechanics and symptom experience. Tests are carried out at a fixed percentage of the peak work rate achieved at baseline testing. The choice of the percentage of the peak work rate is crucial and higher percentages lead to shorter tests at baseline. Lower relative work rates lead to long tests or even "infinite" tests if the work rate is below the critical power [35]. The chosen relative work rate also impacts on the effect size of interventions due to the shape of the power–duration curve (fig. 2a).

In clinical trials, most studies used a work rate of 75% of peak work rate. One study applied a work rate of 75% of peak but if the exercise duration was more than 15 min, work rate was increased to 90% of peak work rate [50]. Few studies used 80% of peak work rate and one study used 50% of peak work. 15 studies looked at several pharmacological interventions and assessed the effects on cycling exercise endurance at one or several time points. A summary of those study groups where patients were tested at a work rate of 75% of peak is provided in figure 2b. With the exception of one intervention, most studies have reported nominal and/or statistically significant improvements in exercise endurance, although there appears to be no relationship between the effects of pharmacotherapy on forced expiratory volume in 1 s (FEV_1) and whole-body endurance (fig. 2c). Whole-body endurance is also affected by skeletal muscle function, which is not improved by pharmacotherapy. This is particularly true for cycling exercises. Therefore, it can be doubted whether more potent bronchodilators will be able to show superiority in terms of enhanced endurance capacity, unless these interventions are combined with improvement of skeletal muscle function. This was shown in one study, where the addition of tiotropium to salmeterol (dual bronchodilation) improved lung function more than one bronchodilator (salmeterol) in combination with fluticasone (71-mL improvement in FEV_1) but did not significantly enhance endurance capacity (15 s) [50].

Endurance walk test

In recent years, walk tests have become more popular to assess the effects of bronchodilators in clinical trials, as it was shown that walk tests may be less impaired by peripheral muscle function

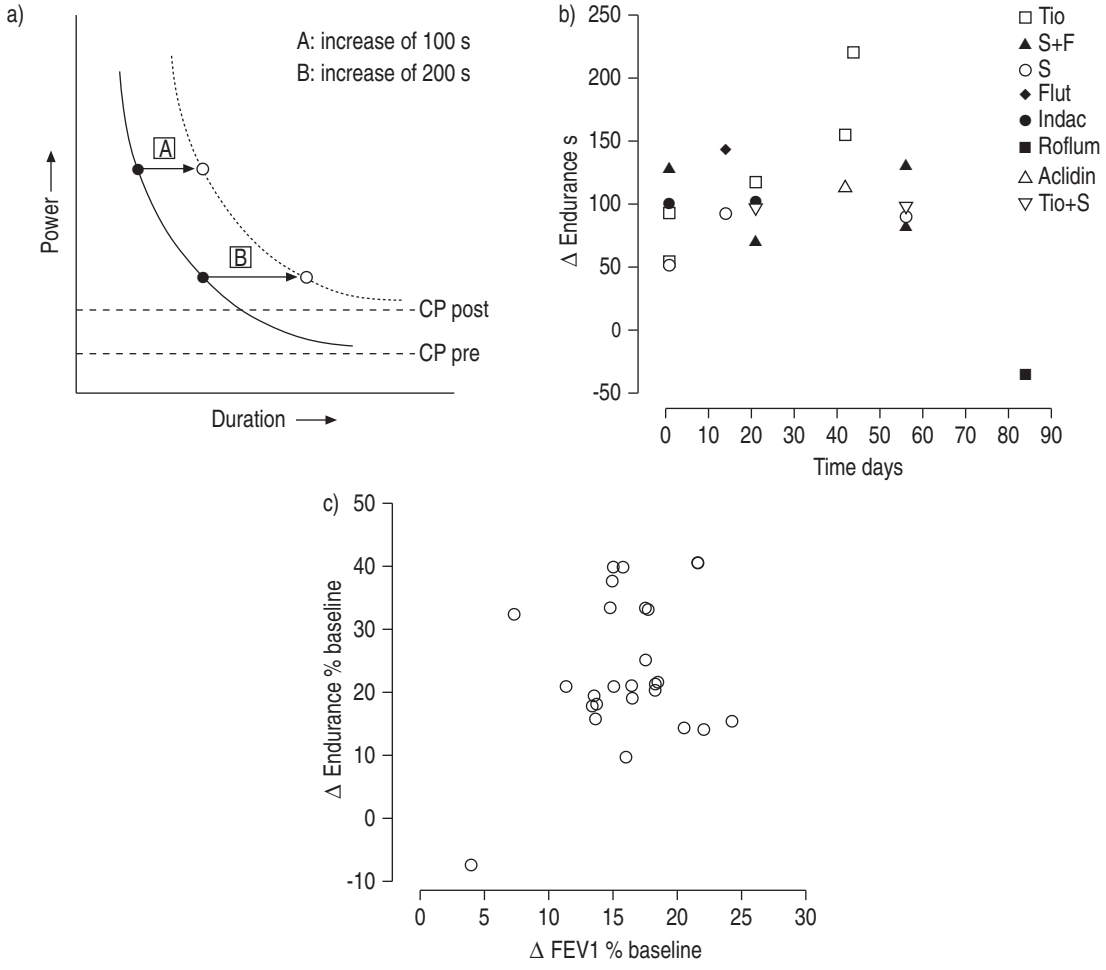


Figure 2. a) Theoretical impact of altered relative work rate or power on exercise duration before and after an intervention, which impacts on the work rate that can be maintained (critical power (CP)). Reproduced and modified from [28] with permission from the author. b) The effect of different bronchodilators on endurance capacity tested in cycling tests at 75% of peak performance at different time-points. Each symbol represents the net effect in one study group compared with the respective control groups in the study. Longer treatment periods (up to 2 months) seem to result in larger effects for a given bronchodilator. Tio: tiotropium; S+F: salmeterol plus fluticasone; S: salmeterol; Flut: fluticasone; Indac: indacaterol; Roflum: roflumilast; Aclidin: aclidinium bromide; Tio+S: tiotropium plus salmeterol. Data from [21, 36–46]. c) Changes in endurance time (% baseline) in relation to changes in lung function (forced expiratory volume in 1 s (FEV_1) % baseline) in studies reporting both patients on treatment and a placebo group. Plotting absolute differences also showed no relationship (not shown). Data from [21, 36–39, 41–44, 46–49].

and, therefore, more responsive to pharmacotherapy. Endurance walk tests can be performed on a treadmill or using field walk tests (endurance shuttle walk test). The latter is performed at a speed that theoretically corresponds to 85% of the peak $V'O_2$ obtained in a prior incremental shuttle walk test [4]. The test can detect the response to bronchodilator therapy in patients with moderate COPD [51], but not in patients with mild COPD [52]. A study using a long-acting anticholinergic in patients with moderate COPD showed an improvement in walking endurance of ~16% after 1 day and ~30% after 3 weeks of treatment. A controlled study using pulmonary rehabilitation as an intervention showed an 18% increase in endurance time in the control group and 63% increase in the rehabilitation group. A significant learning effect has been reported in the endurance shuttle walk test but after a practice walk test the test can be conducted with good reliability. Recently, a protocol has been proposed to test whole-body endurance on a treadmill at 90% of peak work rate

obtained in a maximal incremental treadmill test [53]. 90% of peak resulted in shorter (6.4 min) endurance tests with more maximal ventilation compared with tests carried out at 80% (11.9 min). Hence, 90% of peak was used in a 2-year multicentre trial. It is of note that treadmills also need to be calibrated. Detailed procedures are available [54].

6-min walk test

The 6MWT is a robust estimate of functional exercise capacity. The test can serve as a marker of disease severity and, as such, has been integrated into the BODE (body mass index, airflow obstruction, dyspnoea, exercise capacity) multidimensional severity classification [55]. Deterioration in 6-min walking distance (6MWD) is a bad prognostic sign [56]. The 6MWT has been used as a ''classical'' outcome of pulmonary rehabilitation. Improvement has been reported to be around 50 m in a meta-analysis of pulmonary rehabilitation [57]. 6MWD is not very responsive to pharmacological interventions, probably as patients do not spontaneously change their walking speed after interventions that enhance lung function only. The reproducibility of the test is an asset in longer-term studies but, so far, no pharmacological studies have looked at the effect of pharmacotherapy on decline in 6MWD. Standardisation of the test is crucial and aspects of standardisation include the instructions given to the patient, the track length and course layout (straight, oval or with corners), and especially encouragement during the test. If patients need to rest during the 6 min, this resting time is not compensated by longer walking time. Although the 6MWT does not elicit significant quadriceps contractile fatigue [16], a recent study reported electrographic evidence of lower median frequency of the spectrum of the electromyogram, suggestive of fatiguing effort in the quadriceps muscle in healthy controls as well as in COPD. A standard operating procedure was published by the ATS [5] and will be updated shortly in an ERS/ATS technical document. The minimally important difference has been studied by several authors and is now estimated, based on triangulation of several methods, to be around 25 to 30 m [29]. This value has been reproduced in several studies using anchor or statistical deduction. When peer comparison is used, 54 m has been proposed [58].

Conclusion

Exercise tests are becoming ''standard'' outcome measures of clinical trials in respiratory disease. Several tests are possible but whole-body endurance tests (cycling, treadmill or endurance shuttle walk test) are the most responsive tests to interventions in patients with COPD. The 6MWT is a robust outcome that is responsive to pulmonary rehabilitation but less so after bronchodilator therapy in COPD.

Statement of Interest

T. Trooster's institution has received honoraria for the author to be a speaker and/or member of advisory boards from Boehringer Ingelheim, Novartis and Acertys. The author has received grant support from the Flemish Research Foundation (grant G.0871.13).

References

1. Food and Drug Administration Center for Drug Evaluation and Research. Guidance for Industry Chronic Obstructive Pulmonary Disease: Developing Drugs for Treatment Draft guidance 2007. Silver Spring, Food and Drug Administration, 2007.
2. European Medicines Agency Afficacy Working Party. Guideline on clinical investigation of medicinal products in the treatment of chronic obstructive pulmonary disease (COPD). 2012. London, European Medicines Agency, 2012.
3. Singh SJ, Morgan MD, Scott S, *et al.* Development of a shuttle walking test of disability in patients with chronic airways obstruction. *Thorax* 1992; 47: 1019–1024.
4. Revill SM, Morgan MD, Singh SJ, *et al.* The endurance shuttle walk: a new field test for the assessment of endurance capacity in chronic obstructive pulmonary disease. *Thorax* 1999; 54: 213–222.
5. ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med* 2002; 166: 111–117.

6. Skumlien S, Hagelund T, Bjortuft O, *et al.* A field test of functional status as performance of activities of daily living in COPD patients. *Respir Med* 2006; 100: 316–323.
7. Kato DJ, Rodgers WM, Stickland MK, *et al.* Impact of peak oxygen uptake and muscular fitness on the performance of activities of daily living in patients with chronic obstructive pulmonary disease. *J Cardiopulm Rehabil Prev* 2012; 32: 400–404.
8. Saibene F, Minetti AE. Biomechanical and physiological aspects of legged locomotion in humans. *Eur J Appl Physiol* 2003; 88: 297–316.
9. Man WD, Soliman MG, Gearing J, *et al.* Symptoms and quadriceps fatiguability following walking and cycling in COPD. *Am J Respir Crit Care Med* 2003; 168: 562–567.
10. Palange P, Forte S, Onorati P, *et al.* Ventilatory and metabolic adaptations to walking and cycling in patients with COPD. *J Appl Physiol* 2000; 88: 1715–1720.
11. Mahler DA, Gifford AH, Waterman LA, *et al.* Mechanism of greater oxygen desaturation during walking compared with cycling in patients with COPD. *Chest* 2011; 140: 351–358.
12. Zhang X, Waterman LA, Ward J, *et al.* Advantages of endurance treadmill walking compared with cycling to assess bronchodilator therapy. *Chest* 2010; 137: 1354–1361.
13. Murray JA, Waterman LA, Ward J, *et al.* Perceptual and physiologic responses during treadmill and cycle exercise in patients with COPD. *Chest* 2009; 135: 384–390.
14. Deschenes D, Pepin V, Saey D, *et al.* Locus of symptom limitation and exercise response to bronchodilation in chronic obstructive pulmonary disease. *J Cardiopulm Rehabil Prev* 2008; 28: 208–214.
15. Pepin V, Saey D, Whittom F, *et al.* Walking *versus* cycling: sensitivity to bronchodilation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2005; 172: 1517–1522.
16. Pepin V, Brodeur J, Lacasse Y, *et al.* Six-minute walking *versus* shuttle walking: responsiveness to bronchodilation in chronic obstructive pulmonary disease. *Thorax* 2007; 62: 291–298.
17. Debigare R, Maltais F, Mallet M, *et al.* Influence of work rate incremental rate on the exercise responses in patients with COPD. *Med Sci Sports Exerc* 2000; 32: 1365–1368.
18. Covey MK, Larson JL, Alex CG, *et al.* Test–retest reliability of symptom-limited cycle ergometer tests in patients with chronic obstructive pulmonary disease. *Nurs Res* 1999; 48: 9–19.
19. Hansen JE, Sun XG, Yasunobu Y, *et al.* Reproducibility of cardiopulmonary exercise measurements in patients with pulmonary arterial hypertension. *Chest* 2004; 126: 816–824.
20. Dyer CA, Singh SJ, Stockley RA, *et al.* The incremental shuttle walking test in elderly people with chronic airflow limitation. *Thorax* 2002; 57: 34–38.
21. Oga T, Nishimura K, Tsukino M, *et al.* The effects of oxitropium bromide on exercise performance in patients with stable chronic obstructive pulmonary disease. A comparison of three different exercise tests. *Am J Respir Crit Care Med* 2000; 161: 1897–1901.
22. Troosters T, Casaburi R, Gosselink R, *et al.* Pulmonary rehabilitation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2005; 172: 19–38.
23. Laveneziana P, Palange P, Ora J, *et al.* Bronchodilator effect on ventilatory, pulmonary gas exchange, and heart rate kinetics during high-intensity exercise in COPD. *Eur J Appl Physiol* 2009; 107: 633–643.
24. Lahaije AJ, van Helvoort HA, Dekhuijzen PN, *et al.* Physiologic limitations during daily life activities in COPD patients. *Respir Med* 2010; 104: 1152–1159.
25. Palange P, Ward SA, Carlsen KH, *et al.* Recommendations on the use of exercise testing in clinical practice. *Eur Respir J* 2007; 29: 185–209.
26. ATS/ACCP Statement on cardiopulmonary exercise testing. *Am J Respir Crit Care Med* 2003; 167: 211–277.
27. Ross RM, Corry DB. Software for interpreting cardiopulmonary exercise tests. *BMC Pulm Med* 2007; 7: 15.
28. Borel B, Provencher S, Saey D, *et al.* Responsiveness of various exercise-testing protocols to therapeutic interventions in COPD. *Pulm Med* 2013; 2013: 410748.
29. Puhan MA, Chandra D, Mosenifar Z, *et al.* The minimal important difference of exercise tests in severe COPD. *Eur Respir J* 2011; 37: 784–790.
30. Revill SM, Morgan MD. Biological quality control for exercise testing. *Thorax* 2000; 55: 63–66.
31. Benzo RP, Paramesh S, Patel SA, *et al.* Optimal protocol selection for cardiopulmonary exercise testing in severe COPD. *Chest* 2007; 132: 1500–1505.
32. Glaser S, Lodziewski S, Koch B, *et al.* Influence of the incremental step size in work rate on exercise response and gas exchange in patients with pulmonary hypertension. *BMC Pulm Med* 2008; 8: 3.
33. Saey D, Debigare R, LeBlanc P, *et al.* Contractile leg fatigue after cycle exercise. A factor limiting exercise in patients with COPD. *Am J Respir Crit Care Med* 2003; 168: 425–430.
34. O'Donnell DE, Travers J, Webb KA, *et al.* Reliability of ventilatory parameters during cycle ergometry in multicentre trials in COPD. *Eur Respir J* 2009; 34: 866–874.
35. Neder JA, Jones PW, Nery LE, *et al.* Determinants of the exercise endurance capacity in patients with chronic obstructive pulmonary disease. The power–duration relationship. *Am J Respir Crit Care Med* 2000; 162: 497–504.
36. O'Donnell DE, Fluge T, Gerken F, *et al.* Effects of tiotropium on lung hyperinflation, dyspnoea and exercise tolerance in COPD. *Eur Respir J* 2004; 23: 832–840.
37. O'Donnell DE, Voduc N, Fitzpatrick M, *et al.* Effect of salmeterol on the ventilatory response to exercise in chronic obstructive pulmonary disease. *Eur Respir J* 2004; 24: 86–94.

38. Maltais F, Hamilton A, Marciniuk D, *et al.* Improvements in symptom-limited exercise performance over 8 h with once-daily tiotropium in patients with COPD. *Chest* 2005; 128: 1168–1178.
39. O'Donnell DE, Sciurba F, Celli B, *et al.* Effect of fluticasone propionate/salmeterol on lung hyperinflation and exercise endurance in COPD. *Chest* 2006; 130: 647–656.
40. Maltais F, Celli B, Casaburi R, *et al.* Aclidinium bromide improves exercise endurance and lung hyperinflation in patients with moderate to severe COPD. *Respir Med* 2011; 105: 580–587.
41. O'Donnell DE, Casaburi R, Vincken W, *et al.* Effect of indacaterol on exercise endurance and lung hyperinflation in COPD. *Respir Med* 2011; 105: 1030–1036.
42. Neder JA, Fuld JP, Overend T, *et al.* Effects of formoterol on exercise tolerance in severely disabled patients with COPD. *Respir Med* 2007; 101: 2056–2064.
43. Travers J, Laveneziana P, Webb KA, *et al.* Effect of tiotropium bromide on the cardiovascular response to exercise in COPD. *Respir Med* 2007; 101: 2017–2024.
44. Magnussen H, Paggiaro P, Schmidt H, *et al.* Effect of combination treatment on lung volumes and exercise endurance time in COPD. *Respir Med* 2012; 106: 1413–1420.
45. O'Donnell DE, Bredenbroker D, Brose M, *et al.* Physiological effects of roflumilast at rest and during exercise in COPD. *Eur Respir J* 2012; 39: 1104–1112.
46. Guenette JA, Raghavan N, Harris-McAllister V, *et al.* Effect of adjunct fluticasone propionate on airway physiology during rest and exercise in COPD. *Respir Med* 2011; 105: 1836–1845.
47. Oga T, Nishimura K, Tsukino M, *et al.* A comparison of the effects of salbutamol and ipratropium bromide on exercise endurance in patients with COPD. *Chest* 2003; 123: 1810–1816.
48. Beeh KM, Wagner F, Khindri S, *et al.* Effect of indacaterol on dynamic lung hyperinflation and breathlessness in hyperinflated patients with COPD. *COPD* 2011; 8: 340–345.
49. van der Vaart H, Postma DS, Grevink R, *et al.* Bronchodilation improves endurance but not muscular efficiency in chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis* 2011; 6: 229–235.
50. Magnussen H, Paggiaro P, Schmidt H, *et al.* Effect of combination treatment on lung volumes and exercise endurance time in COPD. *Respir Med* 2012; 106: 1413–1420.
51. Bedard ME, Brouillard C, Pepin V, *et al.* Tiotropium improves walking endurance in COPD. *Eur Respir J* 2012; 39: 265–271.
52. Gagnon P, Saey D, Provencher S, *et al.* Walking exercise response to bronchodilation in mild COPD: a randomized trial. *Respir Med* 2012; 106: 1695–1705.
53. Cooper CB, Abrazado M, Legg D, *et al.* Development and implementation of treadmill exercise testing protocols in COPD. *Int J Chron Obstruct Pulmon Dis* 2010; 5: 375–385.
54. Bishop PA, Bamman MM, Fortney S, *et al.* Procedures for Exercise Physiology Laboratories. Hanover, NASA, 1998.
55. Celli BR, Cote CG, Marin JM, *et al.* The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. *N Engl J Med* 2004; 350: 1005–1012.
56. Polkey MI, Spruit MA, Edwards LD, *et al.* Six-minute-walk test in chronic obstructive pulmonary disease: minimal clinically important difference for death or hospitalization. *Am J Respir Crit Care Med* 2013; 187: 382–386.
57. Lacasse Y, Goldstein R, Lasserson TJ, *et al.* Pulmonary rehabilitation for chronic obstructive pulmonary disease. *Cochrane Database Syst Rev* 2006; CD003793.
58. Redelmeier DA, Bayoumi AM, Goldstein RS, *et al.* Interpreting small differences in functional status: the six minute walk test in chronic lung disease patients. *Am J Respir Crit Care Med* 1997; 155: 1278–1282.

Chapter 14

Omics

Daniel J. Kass and Naftali Kaminski#*

SUMMARY: The promise of personalised medicine is to employ the most basic molecular make-up of a patient to drive individualised, efficacious and cost-effective treatments. Omics, technologies that measure biological function on a global basis, is the key to opening the door to personalised medicine. However, advances in omics technologies have not been incorporated into current clinical trial design in pulmonary medicine. We suggest that the next generation of treatments will be driven by the collection of omics data in patients participating in pulmonary clinical trials. This is critically important because patient populations presenting with chronic advanced lung diseases such as asthma, chronic obstructive pulmonary disease or idiopathic pulmonary fibrosis are heterogeneous and complex, and thus to prove outcomes, studies need to be large and costly. Integrating omics in our clinical trial designs will allow trialists to focus treatments on patients likely to respond and to identify clinically relevant surrogate outcome and drug effect biomarkers. Omics will also lead to unprecedented opportunities for biological discovery in the pathogenesis of lung disease. In this chapter, we focus on how omics technologies will address the shortcomings of current clinical trial designs and what the challenges of incorporating omics into clinical trials are.

*Dorothy P. and Richard P. Simmons Center for Interstitial Lung Disease and the Division of Pulmonary, Allergy, and Critical Care Medicine, University of Pittsburgh, PA, and #Pulmonary, Critical Care, and Sleep Medicine, Yale School of Medicine, Yale University, New Haven, CT, USA.

Correspondence: D.J. Kass, University of Pittsburgh, 3459 Fifth Ave, MUH 628NW, Pittsburgh, PA 15213, USA.
Email: kassd2@upmc.edu

Eur Respir Monogr 2013; 62: 179–187.
Copyright ERS 2013.
DOI: 10.1183/1025448x.10045712
Print ISBN: 978-1-84984-044-6
Online ISBN: 978-1-84984-045-3
Print ISSN: 1025-448x
Online ISSN: 2075-6674

A major obstacle faced by healthcare providers and researchers in lung disease is how to account for the incredible diversity of patients who are characterised by the current definitions of lung disease. Chronic obstructive pulmonary disease (COPD), the third leading cause of death in the USA [1] is a disease characterised by extraordinary heterogeneity [2, 3]. A first year medical student learns to distinguish the classical images of the "blue bloater" from the "pink puffer" [4] and can identify the diverse aetiologies of COPD ranging from smoking to α_1-antitrypsin (α_1-AT) deficiency. Nevertheless, while these patients are clearly very different, they are grouped together in terms of diagnosis and treatment as COPD. Similar controversies exist in other lung diseases such as asthma and interstitial lung disease (ILD), for example. With the proliferation of more sensitive radiographic and molecular technologies, a more thorough characterisation of the tremendous heterogeneity of lung disease appears to be within our grasp. This imperative underlies the clarion call of the National Research Council of the US National Academy of Sciences to devise a new taxonomy of human disease based on molecular biology [5].

How is disease heterogeneity currently incorporated into clinical trial design? The current paradigm for clinical trial design groups patients by overt signs and symptoms but largely ignores

the more subtle but critical concepts of ''subphenotype'' and ''endophenotype''. Subphenotype is a term that is employed widely in the pulmonary literature to describe the heterogeneous behaviours of disease in different patients [6–11]. Endophenotype refers to a heritable and quantifiable trait that is associated with illness in a population but is primarily state independent; that is, it manifests itself in an individual whether or not the disease is active [12]. Those who conduct clinical trials do attempt ''enrichment'' procedures to identify the patients with the subphenotype in whom the particular drug or intervention will most likely generate a signal [13, 14]. However, these efforts are usually crude because they are based on the external phenotypic manifestations of the disease instead of the biological mechanisms that lead to disease. The dependence on disease manifestations instead of molecular mechanisms in patient enrichment often ends up in one of two unwanted outcomes: dilution of effect or over-fitting. Most commonly, because of sample size and statistical power considerations, grouping of disparate patients leads to dilution of results and inability to identify populations of responders. In some other cases, efforts to identify ''pure'' populations miss the common underlying mechanistic endophenotype and leads to over-fitting of patient populations and to limited relevance of the results.

We suggest that the solution to this problem is the routine use of omics data, new technologies that permit measurement of biological activity on a global basis [15], which may provide the resolution necessary to distinguish patients from one another by uncovering subphenotypes and endophenotypes.

Continued adherence to clinical trial designs that ignore the advances in omics, we suggest, will doom many potentially efficacious medications to failure. After years of costly clinical trials, we have few effective evidence-based treatments. The incorporation of omics technologies (table 1) into clinical trials in pulmonary medicine will directly redress our collective failures to uncover new treatments.

In this chapter, we will discuss the benefits and the challenges of integrating omics modalities into clinical trials in lung disease.

The benefits

Why study omics data in clinical trials?

The integration of omics data collection in clinical trials is well underway in oncology [14] but to date has failed to penetrate many other medical disciplines with the same enthusiasm.

Omics data will allow the pulmonary research community to: 1) examine the global shifts of known disease-relevant mechanisms; 2) discover previously unsuspected mechanisms of disease; 3) provide a biological explanation for the ''unexplained'', such as unpredictable adverse effects, or

Table 1. Omics technologies

Genomics	Study of the complete sequence of DNA in a cell or organism. Genomics analysis can detect single nucleotide polymorphisms, insertions, deletions, copy number variation
Transcriptomics	Sometimes categorised under genomics, transcriptomics is the study of the complete set of RNA transcripts from DNA in a cell or tissue. The two dominant technologies are microarrays and RNA sequencing (RNAseq)
Epigenomics	The study of chemical modifications to DNA or to the histones that bind to DNA. This includes DNA methylation
Proteomics	The study of the complete set of proteins expressed by a cell, tissue or organism
Metabolomics	Study of the complete set of small molecule metabolites including carbohydrates, lipids, *etc.*

Definitions are based on [16].

why there are varying responses to drugs; and 4) identify spurious differences in response between drug and placebo due to different population stratification.

While omics data have significantly impacted our basic understanding of disease processes in the lung, clinical trials in lung disease (with the exception of lung cancer) have largely ignored omics approaches to date. We suggest that omics approaches to respiratory disease should be employed now because they are feasible and will lead to unprecedented clinical and biological discovery.

Feasibility

The first argument in favour of incorporating omics data into clinical trial design may be a utilitarian one: because we can. This view is not shared by many who suggest that the cost of omics data collection in clinical trials is exorbitant and the technology necessary to integrate these data is not quite ready for clinical practice [17]. We anticipate that the growth of personalised medicine strategies will demand routine collection of omics data during clinical trials. Patients in a clinical trial represent a "captive audience." The collection of an additional 50 mL of blood for isolation of peripheral blood mononuclear cells (PBMC) for omic analysis could easily be integrated into a research protocol. For those patients who require a biopsy for diagnosis, such as sarcoidosis patients, omic analysis of lung biopsies may become a routine entry test into clinical trials. Point-of-care and cheaper methods of genomic analysis, for example, are already available and may allow even longitudinal collection of data. The "$1000 genome" may be within reach [18]: the US National Human Genome Research Institute (NHGRI) quotes a cost of less than US$10 000 for sequencing one human genome [19]. These data are probably outdated. Biotechnology companies have announced whole genome [20] or exome [21] sequencing costing US$1000 and genome-scale transcriptomic profiling costs less than US$300. This will permit potential identification of biomarkers and gene–drug interactions. Because the data can be obtained, we should not squander the opportunities. The large scale collection of omics data may be useful even if the intervention (*e.g.* drug) fails to show a difference. Imagine the potential of genomic analysis of other seminal clinical trials in medicine. It is tantalising to consider how much clinically important data can come from the integration of omics data collection into the clinical studies of patients where greater than 1000 or even 10 000 patients are enrolled.

Clinical discovery: biomarkers

One goal of integrating omics data collection into clinical trials is to identify clinically relevant biomarkers. Biomarkers are measurable characteristics that can identify a physiological or pathological state of an organism [22]. An outstanding review by SIMON and ROYCHOWDHURY [23] suggested that biomarker discovery from genomic data in cancer clinical trials can be predictive, which drug or combination drugs identified at the point of care will work. A predictive biomarker may also identify a patient at the greatest risk for adverse effects. Perhaps the best example of a predictive biomarker in non-malignant lung disease is the G551D mutation in the cystic fibrosis transmembrane conductance regulator (CFTR), which is used to identify patients who will respond to the newly-approved CFTR-modulating drug ivacaftor [24, 25]. Patients with asthma exhibiting a T-helper cell (Th) type 2 gene expression profile are more likely to respond to corticosteroid treatment compared to non-Th2 patients [10]. A biomarker may also be prognostic, identifying who will need treatment or whose disease will be quiescent. When used longitudinally, such a biomarker may grade the patient's progress through a course of treatment. RICHARDS *et al.* [26] reported that plasma analysis of five proteins in patients with IPF by a high-throughput protein-detection technique predicted prognosis in 241 patients with IPF. The potential application of such biomarkers, if validated prospectively, is to prioritise patients for treatment (if available) or for transplant. Biomarkers may also provide pharmacogenomic information, identifying which dose of a medication will be effective in a particular patient. For example, a patient's genotype for CYP2C9 and VKORC1 may explain between 30 and 40% of the total variation of the final warfarin dose (reviewed in [27]). Thus integration of omics technologies in clinical studies will allow both the measurement of currently recognised markers and the discovery of new and more accurate biomarkers.

Biological discovery

"Bedside-to-bench" research will benefit tremendously from the collection of human omics data in clinical trials. In our laboratory, we have focused our efforts on generating a large repository of clinical and genomic data on all forms of diffuse parenchymal lung disease based on analysis of carefully phenotyped patient samples and local expertise in genomics, systems biology and pulmonary medicine and biology. This led in IPF, for example, to identification of a large number of genes that have been studied employing *in vivo* and *in vitro* models of fibrosis and contributed a wealth of information on the pathogenesis of disease [28–38]. Naturally, while studies in animal models of disease provide mechanistic insights, they do not provide an accurate representation of the human disease. In contrast, a clinical trial that utilises a mechanistic perturbation to modify disease represents an opportunity to study the biological effects of this perturbation, even if the drug fails to achieve the primary effect. Thus integrating omics analysis, especially transcriptomic, proteomic and metabolomic, during the course of a drug study will generate unprecedented insights into disease and pharmacological mechanisms. The sharing of such results in a well-curated and publically available lung disease-targeted database will provide scientists with unique opportunities for understanding the biology of disease and the effects of drugs.

The challenges

Whom to study, what data to analyse and how?

Organ

A challenge facing those designing trials is to decide what will be analysed by omics technology. Since oncology is driven by a biopsy of the tumour, then a global, omics-based analysis of the tumour is the natural starting point. In non-malignant lung disease, frequently diagnoses are made by clinical criteria without a biopsy. In only a few cases are diagnoses made by tissue biopsy as in the case of ILD or sarcoidosis. Sampling of multiple compartments has been advocated for chronic lung disease: exhaled breath, induced sputum, bronchial brushings, bronchoalveolar lavage (BAL) (fluid and cells) and blood (serum, plasma or cells). The choice about what to analyse is not often clear and depends more on the expertise of the investigators and the perceived relevance to the disease [11] then on hard data. For example, bronchial brushings may be more relevant to an asthma study while peripheral blood may make more sense for a sarcoidosis study. Unfortunately, at this stage, beyond intuition, we lack the data that will allow us to make such a choice. However, it is obvious that, in most cases, lung biopsies will not be available or justified in chronic non-malignant lung diseases and thus focus on systematic omics analysis of lung "surrogate" tissues and comparison of the utility and information content of different compartments in different lung diseases will be required.

What kind of omics data?

As outlined in table 1, omics analyses study the compendium of changes in a patient's DNA, RNA, protein and metabolites. Technologies are constantly evolving. Investigators will have to decide which particular omics technology to employ, and the considerations will vary based on coverage, resolution, cost and standardisation (as discussed in the next section). Naturally, different levels of omics data provide different types of information and thus will be assayed at different frequencies. For example, whole genome sequencing should be performed once to yield information about genetic predisposition. However, transcriptomics that are indicative of changes in whole genome gene expression will likely require frequent assessment, before and after treatment, and the same for proteomic assessment. It is unclear how frequently changes in the epigenome need to be studied. Because of the lack of information on these topics, small observational studies that assess the temporal change in the patient's transcriptome, proteome, epigenome and metabolome in different surrogate compartments are required.

Standardisation

The choice of what to study and what modality to employ goes to the critical question of standardisation. Laboratories that generate clinical omics data upon which clinical decisions are made need to consider standardised laboratory techniques to ensure generalisable solutions for patients across the country [23]. In the USA, the Centers for Medicare and Medicaid Services (CMS) regulate all laboratory testing (except research) through the Clinical Laboratory Improvements Amendment (CLIA) [39]. As noted, standardisation applies to the training of personnel, the equipment and reagents used, quality control, and interpretation of results [23]. While these are research tools at the beginning, the translation of these results to clinical practice will require standardisation and transparency in terms of methodology. Applying very rigorous standards of evidence for the development of omics-based testing prior to their use in clinical trials was the goal of the Institute of Medicine's report [40].

How will the data be organised and shared?

Because of how large the datasets can become, there have been discussions of how results should be reported [15, 41]. A number of databases have arisen to manage the rapid and voluminous accumulation of genomics data [16, 23], however, the majority of them are not easy to use, do not address the specific challenges of lung disease, do not allow for easy access to multi-hierarchical omics data and do not contain detailed phenomic information. In pulmonary medicine, the lung genomics research consortium (LGRC) was funded to be the largest research database of clinical and genomics information in lung disease [11] and provides unprecedented access to parallel phenomic, transcriptomic, genomic and epigenomic data, but it is unclear whether this effort will be sustained. As these databases proliferate, with advanced algorithms for analysing data, there will be issues of standardisation and interaction between the databases. The gathering of high throughput omics data will also invariably raise issues regarding how data will be stored and protected. The recent failures of banks to protect both money and information from cybercrime suggest that the barriers to potentially sensitive and private data can be easily compromised. In order for these databases to be clinically useful, they will have to be integrated so that when these biomarkers come to practice, clinicians will not have to sift through and interpret diverse sources of information [23]. In oncology, where the omics discoveries outpace those in other fields, clinicians must keep up with the constantly growing list of omics targets [23]. To address this problem, online clinical decision tools are being developed, and some institutions have developed multidisciplinary rounds including experts in cancer genomics and bioinformatics to help guide clinical decision making (summarised in [23]).

How will clinical trials be designed?

In oncology, paradigm-shifting trial design based on omics technology is transforming traditional concepts of clinical trials. The idea of histology-independent and mutation-specific stratification of patients for clinical trial design in oncology of patients rather than by histology is gaining traction [14]. The rationale behind this model is that a drug that targets a particular mutation is more likely to show an effect if the clinical trial population is enriched with patients who share this mutation rather than an unselected population based upon histology. The new paradigm suggests that current models for clinical trial design miss potentially efficacious medications since the stratification of patients does not account for subphenotypes or endophenotypes. Consider the hypothetical example shown in figure 1. A large population with idiopathic pulmonary fibrosis (IPF) is randomised to drug or placebo in a double-blinded fashion. In addition to standard outcomes such as mortality, change in pulmonary function and quality of life (QoL), PBMCs are collected from each patient at the beginning and at the end of the trial. RNA is isolated from the PBMCs and subjected to gene expression profiling by microarray analysis. As shown, mean forced vital capacity (FVC) at the conclusion of the trial is not significantly different compared to placebo. However, the research team notes on *post hoc* analysis that there are groups of patients with distinct patterns of biomarkers that cluster together at the far ends of the distribution. One group of patients whose pulmonary function greatly improved shared the GG genotype of the

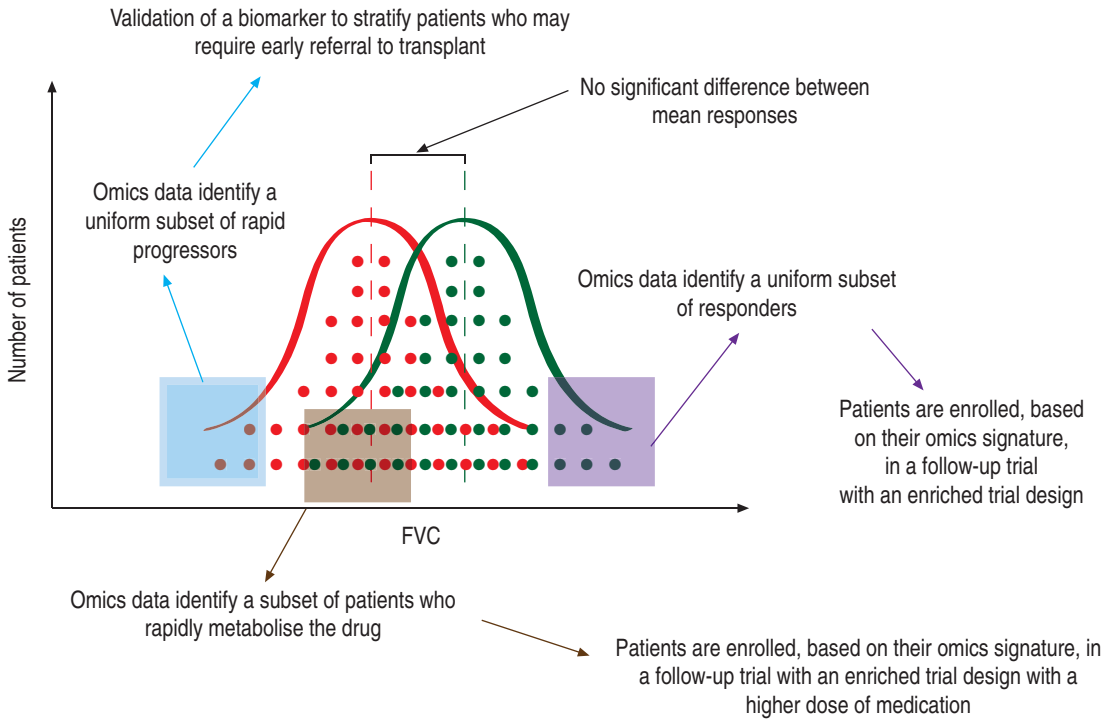

Validation of a biomarker to stratify patients who may require early referral to transplant
No significant difference between mean responses
Omics data identify a uniform subset of rapid progressors
Omics data identify a uniform subset of responders
Patients are enrolled, based on their omics signature, in a follow-up trial with an enriched trial design
Number of patients
FVC
Omics data identify a subset of patients who rapidly metabolise the drug
Patients are enrolled, based on their omics signature, in a follow-up trial with an enriched trial design with a higher dose of medication

Figure 1. Omics data may lead to clinically relevant biomarkers. The results of a hypothetical, randomised, double-blinded placebo control trial in idiopathic pulmonary fibrosis are shown. Mean changes in forced vital capacity (FVC) are not significantly different between placebo (red) and drug (green). However, *post hoc* analysis of omics data identify patients who may respond positively to the drug (purple), patients who rapidly clear the drug (brown) and patients who rapidly progress (blue). Follow-up trials are proposed to test the validity of these biomarkers.

rs35705950 SNP in the MUC5B promoter [42], for example, suggesting that the drug was effective in this subpopulation. Another group of patients was identified as rapid progressors by sharing the deleterious rs6743890 TOLLIP SNP [43]. We argue that the identification of these clusters with certain clinical outcomes would not be possible without an omics approach to clinical trials. Perhaps there is a small group of IPF patients with evidence of immune activation [44, 45] who may have benefitted from corticosteroids as part of the PANTHER trial [46]. Similarly, biomarkers in asthma [47–49] and COPD [50–52] may identify a highly specified group of patients who will benefit not just from the next generation of treatments but possibly even older medications that were shelved because they failed in highly diluted populations.

Trialists will have to decide what their approach will be to biomarker discovery. Will they employ a supervised method in which clinical outcomes are related directly to genomic methods, for example, *versus* an unsupervised method in which omics data are clustered together and then assessed for clinical effects [15]? As investigators employ more sophisticated subphenotype- and endophenotype-driven enrichment techniques, they will struggle with enrolment as the enriched population will be smaller [14]. The identification of this smaller population will be essential since it is likely that a study would fail in the presence of an unselected population [14].

Implementation of omics technologies in clinical research will require training of a whole new generation of investigators and clinicians. In our view, this is the most glaring obstacle in implementing omics approaches in clinical pulmonary research. While computational biologists, geneticists, experts in omics, statisticians, molecular biologists and pulmonary clinical researchers are trained in abundance, the key component required to complement this diverse interdisciplinary approach is the absence of pulmonary researchers and physicians who are well trained in omics disciplines. We have previously advocated the expansion of mechanisms for

training future leaders in omics in lung disease [11] and we strongly believe that resources need to be allocated for training pulmonary clinicians and investigators in omics disciplines. This could be for example by including omics technologies in fellowship training, addition of Masters and certificate programmes similar to those in public health in medical omics, and enrichment of available programs in public health and translational medicine with courses in omics literacy.

Conclusion

It has been noted that the progress of translating omics-based testing into direct clinical applications has been slowed by expense, inconsistent standards and lack of appropriate study designs [40]. Most of the studies have been performed or are planned in oncology. Omics technology has had an impact on the understanding of lung disease, but the incorporation of omics into pulmonary clinical trials is very much in its infancy. Because these modalities are new, we should not excuse avoidance of the omics approach altogether. The pulmonary community should embrace the challenges and strive to design paradigm-shifting clinical trials. This idea leaves open the potentially heretical possibility that COPD and IPF patients, for example, will be included in the same trial because omics data suggest that they share similar subphenotypes and endophenotypes. We advocate that omics data be collected in all clinical trials and suggest that the government agencies that fund research worldwide and the pharmaceutical industry will need to furnish the resources necessary to train the next generation of trialists with knowledge to design the next generation of clinical trials. These investigators should have all the resources necessary to collect the critical omics data. Only then will we truly see the promises of personalised pulmonary medicine. We cannot (and should not) afford more failed clinical trials.

Support Statement

This manuscript was supported by NIH grants HL083085 (D.J. Kass) and HL108642, HL112707, HL095397 (N. Kaminski).

Statement of Interest

D.J. Kass is currently a sub-investigator on multiple clinical trials in pulmonary fibrosis sponsored by Arresto Biosciences, Gilead, Intermune, Stromedix, Boehringer-Ingelheim and Pfizer, but has no role in the planning or design of these clinical trials. N. Kaminski has received money for consultancy from Takeda, Vertex, Sanofti, Biogen, InterMune, Bayer and Promediro, payment for lectures from Air Foundation, and his institution has received grants from Gilead and Celgene, and holds a patent related to peripheral blood markers.

References

1. Minino AM, Murphy SL, Xu J, *et al.* Deaths: final data for 2008. *Natl Vital Stat Rep* 2011; 59: 1–126.
2. Zeskind JE, Lenburg ME, Spira A. Translating the COPD transcriptome: insights into pathogenesis and tools for clinical management. *Proc Am Thorac Soc* 2008; 5: 834–841.
3. Wedzicha JA. The heterogeneity of chronic obstructive pulmonary disease. *Thorax* 2000; 55: 631–632.
4. Robinson RW, White DP, Zwillich CW. Relationship of respiratory drives to dyspnea and exercise performance in chronic obstructive pulmonary disease. *Am Rev Respir Dis* 1987; 136: 1084–1090.
5. Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease. Washington, National Academies Press, 2011.
6. Green RH, Brightling CE, Bradding P. The reclassification of asthma based on subphenotypes. *Curr Opin Allergy Clin Immunol* 2007; 7: 43–50.
7. Hunt J. If it smells like a duck, it might be an asthma subphenotype. *Am J Respir Crit Care Med* 2007; 175: 975–976.
8. Snoeck-Stroband JB, Lapperre TS, Gosman MM, *et al.* Chronic bronchitis sub-phenotype within COPD: inflammation in sputum and biopsies. *Eur Respir J* 2008; 31: 70–77.
9. Chandra D, Sciurba FC, Gladwin MT. Endothelial chronic destructive pulmonary disease (E-CDPD): is endothelial apoptosis a subphenotype or prequel to COPD? *Am J Respir Crit Care Med* 2011; 184: 153–155.
10. Woodruff PG, Modrek B, Choy DF, *et al.* T-helper type 2-driven inflammation defines major subphenotypes of asthma. *Am J Respir Crit Care Med* 2009; 180: 388–395.

11. Kass DJ, Kaminski N. Evolving genomic approaches to idiopathic pulmonary fibrosis: moving beyond genes. *Clin Transl Sci* 2011; 4: 372–379.
12. Gottesman II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. *Am J Psychiatry* 2003; 160: 636–645.
13. US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research, and Center for Devices and Radiological Health. Guidance for Industry: Enrichment Strategies for Clinical Trials to Support Approval of Human Drugs and Biologic Products; 2012. www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM332181.pdf Date last updated: December 2012. Date last accessed: June 13, 2013.
14. Sleijfer S, Bogaerts J, Siu LL. Designing transformative clinical trials in the cancer genome era. *J Clin Oncol* 2013; 31: 1834–1841.
15. Ghosh D, Poisson LM. "Omics" data and levels of evidence for biomarker discovery. *Genomics* 2009; 93: 13–16.
16. Ahmed J, Meinel T, Dunkel M, *et al.* CancerResource: a comprehensive database of cancer-relevant proteins and compound interactions supported by experimental knowledge. *Nucleic Acids Res* 2011; 39: D960–D967.
17. Koscielny S. Why most gene expression signatures of tumors have not been useful in the clinic. *Sci Transl Med* 2010; 2: 14ps2.
18. Institute NHGR. Beyond the Beginning: The Future of Genomics. 2001. www.genome.gov/10001294 Date last updated: 1 October, 2012. Date last accessed: June 13, 2013.
19. Wetterstrand K. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program. 2013. www.genome.gov/sequencingcosts. Date last updated: 16 July, 2013. Date last accessed: June 13, 2013.
20. DeFrancesco L. Life Technologies promises $1,000 genome. *Nat Biotechnol* 2012; 30: 126.
21. 23andMe. Exome 80x. 2011. www.23andme.com/exome/ Date last updated: 2013. Date last accessed: June 13, 2013.
22. Wheelock CE, Goss VM, Balgoma D, *et al.* Application of omics technologies to biomarker discovery in inflammatory lung diseases. *Eur Respir J* 2013; 42: 802–825.
23. Simon R, Roychowdhury S. Implementing personalized cancer genomics in clinical trials. *Nat Rev Drug Discov* 2013; 12: 358–369.
24. Davies JC, Wainwright CE, Canny GJ, *et al.* Efficacy and Safety of Ivacaftor in Patients Aged 6 to 11 Years with Cystic Fibrosis with a G551D Mutation. *Am J Respir Crit Care Med* 2013; 187: 1219–1225.
25. Ramsey BW, Davies J, McElvaney NG, *et al.* A CFTR potentiator in patients with cystic fibrosis and the G551D mutation. *N Engl J Med* 2011; 365: 1663–1672.
26. Richards TJ, Kaminski N, Baribaud F, *et al.* Peripheral blood proteins predict mortality in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2012; 185: 67–76.
27. Wang L, McLeod HL, Weinshilboum RM. Genomics and drug response. *N Engl J Med* 2011; 364: 1144–1153.
28. Pandit KV, Corcoran D, Yousef H, *et al.* Inhibition and role of let-7d in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2010; 182: 220–229.
29. Bridges RS, Kass D, Loh K, *et al.* Gene expression profiling of pulmonary fibrosis identifies Twist1 as an antiapoptotic molecular "rectifier" of growth factor signaling. *Am J Pathol* 2009; 175: 2351–2361.
30. Kass DJ, Yu G, Loh KS, *et al.* Cytokine-like factor 1 gene expression is enriched in idiopathic pulmonary fibrosis and drives the accumulation of $CD4^+$ T cells in murine lungs: evidence for an antifibrotic role in bleomycin injury. *Am J Pathol* 2012; 180: 1963–1678.
31. Zuo F, Kaminski N, Eugui E, *et al.* Gene expression analysis reveals matrilysin as a key regulator of pulmonary fibrosis in mice and humans. *Proc Natl Acad Sci USA* 2002; 99: 6292–6297.
32. Rosas IO, Richards TJ, Konishi K, *et al.* MMP1 and MMP7 as potential peripheral blood biomarkers in idiopathic pulmonary fibrosis. *PLoS Med* 2008; 5: e93.
33. Hsu E, Shi H, Jordan RM, *et al.* Lung tissues in patients with systemic sclerosis have gene expression patterns unique to pulmonary fibrosis and pulmonary hypertension. *Arthritis Rheum* 2011; 63: 783–794.
34. Vuga LJ, Ben-Yehudah A, Kovkarova-Naumovski E, *et al.* WNT5A is a regulator of fibroblast proliferation and resistance to apoptosis. *Am J Respir Cell Mol Biol* 2009; 41: 583–589.
35. Armanios MY, Chen JJ, Cogan JD, *et al.* Telomerase mutations in families with idiopathic pulmonary fibrosis. *N Engl J Med* 2007; 356: 1317–1326.
36. Seibold MA, Wise AL, Speer MC, *et al.* A common MUC5B promoter polymorphism and pulmonary fibrosis. *N Engl J Med* 2011; 364: 1503–1512.
37. Zhang Y, Noth I, Garcia JG, *et al.* A variant in the promoter of MUC5B and idiopathic pulmonary fibrosis. *N Engl J Med* 2011; 364: 1576–1577.
38. Lawson WE, Grant SW, Ambrosini V, *et al.* Genetic mutations in surfactant protein C are a rare cause of sporadic cases of IPF. *Thorax* 2004; 59: 977–980.
39. Clinical Laboratory Improvement Amendments. 2013. Available from: www.cms.gov/Regulations-and-Guidance/Legislation/CLIA/index.html?redirect=/clia/ Date last updated: 1 August, 2013. Date last accessed: June 13, 2013.
40. Micheel CM, Nass SJ and Omen GS, eds. Committee on the Review of Omics-Based Tests for Predicting Patient Outcomes in Clinical Trials Board on Health Care Services, Board on Health Sciences Policy. Evolution of Translational Omics: Lessons Learned and the Path Forward. 2012.
41. McShane LM, Altman DG, Sauerbrei W, *et al.* Reporting recommendations for tumor marker prognostic studies. *J Clin Oncol* 2005; 23: 9067–9072.

42. Peljto AL, Zhang Y, Fingerlin TE, *et al.* Association between the MUC5B promoter polymorphism and survival in patients with idiopathic pulmonary fibrosis. *JAMA* 2013; 309: 2232–2239.
43. Noth I, Zhang Y, Ma SF, *et al.* Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: a genome-wide association study. *Lancet Respir Med* 2013; 1: 8.
44. Kahloon RA, Xue J, Bhargava A, *et al.* Patients with idiopathic pulmonary fibrosis with antibodies to heat shock protein 70 have poor prognoses. *Am J Respir Crit Care Med* 2013; 187: 768–775.
45. Wells AU, Kelleher WP. Idiopathic pulmonary fibrosis pathogenesis and novel approaches to immunomodulation: we must not be tyrannized by the PANTHER data. *Am J Respir Crit Care Med* 2013; 187: 677–679.
46. Raghu G, Anstrom KJ, King TE Jr, *et al.* Prednisone, azathioprine, and *N*-acetylcysteine for pulmonary fibrosis. *N Engl J Med* 2012; 366: 1968–1977.
47. Woodruff PG, Boushey HA, Dolganov GM, *et al.* Genome-wide profiling identifies epithelial cell genes associated with asthma and with treatment response to corticosteroids. *Proc Natl Acad Sci USA* 2007; 104: 15858–15863.
48. Ober C, Tan Z, Sun Y, *et al.* Effect of variation in CHI3L1 on serum YKL-40 level, risk of asthma, and lung function. *N Engl J Med* 2008; 358: 1682–1691.
49. Moffatt MF, Kabesch M, Liang L, *et al.* Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* 2007; 448: 470–473.
50. Huang SL, Su CH, Chang SC. Tumor necrosis factor-alpha gene polymorphism in chronic bronchitis. *Am J Respir Crit Care Med* 1997; 156: 1436–1439.
51. Qiu W, Baccarelli A, Carey VJ, *et al.* Variable DNA methylation is associated with chronic obstructive pulmonary disease and lung function. *Am J Respir Crit Care Med* 2012; 185: 373–381.
52. Hunninghake GM, Cho MH, Tesfaigzi Y, *et al.* MMP12, lung function, and COPD in high-risk populations. *N Eng J Med* 2009; 361: 2599–2608.

ERM 62 December 2013
CME Credit Application Form

This book has been accredited by the European Board for Accreditation in Pneumology (EBAP) for 5 CME credits.

To receive CME credits, read this issue of the Monograph, indicate the correct responses to the educational questions and complete the requested information. To return your application, you can either:

> use this form and return it completed to ERS Publications Office, 442 Glossop Road, Sheffield, S10 2PX, UK; fax to +44-114-2665064; or e-mail to info@ersj.org.uk

> fill in the online form available on the table of contents for the issue at erm.ersjournals.com

Certificates will be e-mailed to the address filled in below. Please allow 4 weeks for processing.

Applicant personal details

ERS Membership No. : .....Date of Birth (DD/MM/YYYY): ..... (if known)

Family Name: ..... First Name: .....

Mailing Address: .....

.....

.....

Postal Code: ..... City: ..... Country: .....

Telephone: +..... E-mail: .....

Educational questions

1. Which of the following statements is true?

❑ There is a strong association between lung function and markers of bronchial inflammation in asthma. ❑ Lung function is predictive of steroid responsiveness in obesity-related asthma. ❑ In patients without previous treatment, levels of exhaled nitric oxide fraction (F_{ENO}) <25 ppb are associated with low likelihood of responsiveness to corticosteroids. ❑ The implication of F_{ENO} in tailoring asthma treatment reduces exacerbations in all phenotypes of asthma.

2.Which of the following statements is true?

❑ Assessment of symptoms and assessment of asthma control are the same. ❑ When assessing symptoms in clinical trials only paper-and-pencil diaries should be used. ❑ Assessment of exacerbations is not important when assessing asthma control. ❑ The Asthma Control Test (ACT) and Asthma Control Questionnaire (ACQ) are both composite scores for the assessment of asthma control. ❑ The ACT requires measurement of spirometry.

3. Which of the following statements is false?

❑ Asthma exacerbations are a transient worsening of asthma. ❑ In clinical trials, exacerbations have been described by a wide variety of definitions. ❑ Exacerbations can be defined as events that require urgent action on the part of the patient and physicians to prevent serious outcome. ❑ Treatment with anti-IgE reduces exacerbations in patients with severe allergic asthma. ❑ In clinical trials, exacerbations should be differentiated into mild, moderate and severe.

Continued on next page

4. Which of the following statements is false?
❑ Mortality from asthma in clinical trials is low. ❑ Chronic use of long-acting β_2-agonists without inhaled corticosteroids has been associated with increased mortality in patients with asthma. ❑ Use of long-acting β_2-agonists in combination with an inhaled corticosteroid is associated with increased mortality in asthma patients. ❑ Due to low death rates, mortality is not a feasible end-point in clinical studies of stable asthma patients. ❑ The validity of asthma-specific mortality as an outcome parameter is very much dependent on the accurate determination of the cause of death.

5. A 30-year-old female experiences recurrent episodes of dyspnoea with intensive coughing, despite treatment with high-dose inhaled corticosteroid, long-acting β_2-agonists and nasal steroids daily. She has never smoked and receives three to four courses of prednisolone per year. Physical examination revealed a body mass index of 41 kg·m^{-2} and wheezing, but no other abnormalities. Lung function showed slightly diminished vital capacity and forced expiratory volume in 1 s (FEV_1) with bronchial hyperactivity. Blood tests showed no abnormalities, except for a radioallergosorbent test, which was positive for grass pollen (2+) and house dust mite (1+). Which is the most appropriate next step?
❑ Increase dose of inhaled steroid. ❑ Add a leukotriene receptor antagonist. ❑ Add theophylline. ❑ Assess patient compliance, potential trigger factors and comorbidities. ❑ Trial with anti-TNF-α.

6. Which of the following statements is true? In a clinical trial in patients with asthma:
❑ Bronchial hyperresponsiveness should be measured at all study visits. ❑ Sputum eosinophils should always be assessed. ❑ Airway resistance is an excellent surrogate marker for airway inflammation. ❑ Specific IgE should be assessed at baseline to determine atopic or non-atopic patients with asthma. ❑ There is no need to assess history of exacerbations at baseline.

7. Which statement concerning forced vital capacity (FVC) in relation to idiopathic pulmonary fibrosis (IPF) is true?
❑ FVC is a validated surrogate for mortality in IPF. ❑ In a clinical trial, a reduction by 5% in FVC decline in the active treatment arm is evidence that the therapy will be associated with improved survival. ❑ FVC inevitably declines prior to an acute exacerbation (AE) of IPF. ❑ When assessed categorically, 6-month change in FVC can be used to predict individuals with IPF with a poor prognosis. ❑ FVC change occurs independently of progressive fibroproliferation.

8. Which of the following is true concerning AEs of IPF?
❑ AEs may be caused by infection. ❑ AEs respond rapidly to therapy with corticosteroids. ❑ On biopsy, AEs are characterised by the finding of diffuse alveolar damage on a background of established fibrosis. ❑ AEs are of little relevance in IPF clinical trials. ❑ AEs rarely result in hospitalisation.

9. In relation to trials of novel compounds for IPF, which of the statements concerning mortality is true?
❑ Mortality is the primary end-point of choice for IPF registration studies. ❑ Surrogate measures of mortality have been validated for use in IPF trials. ❑ All-cause mortality in IPF is the only end-point that cannot be confounded by missing values. ❑ Mortality has never been used as a primary clinical trial end-point in IPF. ❑ An IPF mortality trial powered to show a beneficial effect of a novel compound would require >12,500 patient-years to complete.

10. When thinking about the treatment of IPF, the general aim of treatment with compounds tested in clinical trials to date has been to:
❑ Slow or prevent the development of fibrosis. ❑ Restore functioning alveoli. ❑ Improve lung function. ❑ Prevent the development of pulmonary hypertension. ❑ Cause resorption of existing fibrosis.

11. Which of the following is true about secondary outcome measures commonly used in IPF clinical trials?
❑ The 6-min walking distance is highly reproducible in individuals with IPF. ❑ Serial change in serum level of MMP-7 correlates with progression of fibrosis in IPF. ❑ Worsening of fibrosis on computed tomography (CT) scan is easily detected, with good reproducibility scores when scans are assessed by multiple expert radiologists. ❑ A reduction in 6-min walking distance reflects worsening fibrosis. ❑ A 6- or 12-month categorical reduction in diffusion capacity of the lung for carbon monoxide (D_{LCO}) by 15% predicts individuals with a poor prognosis.

12. Proof-of-concept and early-phase studies are an important part of the drug discovery pathway. Which statement is true of such studies in IPF?
❑ FVC is the best primary outcome measure. ❑ Design is informed by animal models of fibrosis. ❑ They are suited to the use of imaging end-points. ❑ They are limited by a lack of short-term, validated efficacy end-points. ❑ They can be effectively powered using panels of plasma biomarkers.

13. Understanding the pathogenesis of IPF is helping to shape the development of novel therapeutic strategies. Which of the following is true?
❑ Fibroblast apoptosis is frequently observed in IPF lungs. ❑ Increased oxidative stress is a hallmark of IPF. ❑ Epithelial damage is rarely observed in IPF. ❑ Changes in the extracellular matrix result in increased elasticity of the fibrotic lung. ❑ Anti-coagulation with warfarin reduces the progression of fibrosis and improves outcomes in IPF.

14. Patient-reported outcome (PRO) measures are increasingly being used in clinical trials. With respect to these measures in IPF, which of the following is true?
❑ The St George's Respiratory Questionnaire (SGRQ) was specifically designed for use in IPF. ❑ The IPF-specific ATAQ (a tool to assess quality of life in IPF) has been validated for clinical trial use. ❑ The minimum clinically important difference in SGRQ has been defined for IPF. ❑ There are no clinically validated cough questionnaires available for use in IPF. ❑ The validity of PRO tools for IPF can be inferred from their use in other respiratory disorders.

15. Which of the following is not a limitation of FEV_1?
❑ Requires cooperation to perform the manoeuvre. ❑ Insensitive to small airways disease. ❑ Coefficient of variation of 25%. ❑ Relative stability in mild lung disease.

16. Which of the following are advantages of the lung clearance index?
❑ More sensitive to mild lung disease than spirometry. ❑ Can be performed during tidal breathing. ❑ Consistent reference ranges across all ages. ❑ All of the above.

17. Which of the following is not true regarding CT scans in cystic fibrosis (CF) patients?
❑ CT scans are responsive to interventions such as dornase alfa and inhaled tobramycin. ❑ CT scans are able to track longitudinal lung disease over time. ❑ Standard operating procedures exist for CT scans in CF patients for multicentre studies. ❑ CT scans appear to be more sensitive to early lung disease than FEV_1.

18. Which of the following is not true regarding bronchoalveolar lavage (BAL) in CF patients?
❑ The accuracy of defining a patient's infectious state, especially in relation to early infection with *Staphylococcus aureus*, has been shown to be superior to other available culture techniques in non-sputum-producing infants and children. ❑ Standard operating procedures do not exist. ❑ Free neutrophil elastase in BAL samples from infants is associated with persistent bronchiectasis. ❑ BAL is the gold standard to assess lower airways inflammation but is invasive because it requires a general anaesthetic.

Continued on next page

19. Which of the following is not true regarding sputum in CF patients?
❑ Approximately 50% of CF patients aged 6–18 years can expectorate. ❑ Higher baseline neutrophil elastase has been associated with more rapid lung function decline. ❑ One of the advantages of the use of sputum as an outcome measure for CF is that it is noninvasive. ❑ Bronchospasm, the main side-effect of sputum induction, occurs in about 25% of CF patients.

20. Which of the following is true regarding intestinal current measurement?
❑ Able to discriminate between healthy patients and patients with CF. ❑ Able to discriminate between pancreatic sufficiency and pancreatic insufficiency. ❑ Promising biomarker because it can directly assess the effect of CFTR correctors. ❑ All of the above.

21. Which of the following is not true regarding pulmonary exacerbations?
❑ It is a surrogate outcome measure because of the association with survival. ❑ It is a risk factor for rate of decline of FEV_1. ❑ There is an agreed definition for mild exacerbations that has been validated. ❑ There is no agreed definition for moderate and severe exacerbations.

22. Mortality is usually the most suitable outcome for randomised controlled trials in critical illness.
❑ True. ❑ False.

23. Outcomes for randomised controlled trials should be selected according to:
❑ The study question. ❑ The phase of the trial. ❑ Trial resource. ❑ All of the above.

24. Regarding pilot trials, outcomes should:
❑ Be mechanistic to examine proof of principle. ❑ Focus on feasibility. ❑ Suit the objective of the pilot trial.

25. Adjudication is best suited for outcomes that are:
❑ Objective but infrequent. ❑ Objective but frequent. ❑ Somewhat subjective and infrequent. ❑ Somewhat subjective and frequent. ❑ Somewhat subjective regardless of frequency.

26. Quality of life measures are:
❑ Patient important. ❑ Easy to measure. ❑ Relatively inexpensive to measure compared to mortality. ❑ Decreasingly used in critical care.

27. What is the definition of physical activity?
❑ Exercise that is planned, structured and repetitive. ❑ The sum of cardiorespiratory endurance, muscular endurance, muscular strength, body composition and flexibility. ❑ Physical fitness combined with exercise capacity. ❑ Any bodily movement produced by skeletal muscles that results in energy expenditure. ❑ A set of basic, everyday tasks related to personal self-care and independent living.

28. Current health recommendations advocate:
❑ 5 min of vigorous activity per day on 5 days a week. ❑ 30 min of moderate physical activity on 5 days a week. ❑ 15 min of swimming or cycling per day on 5 days a week. ❑ 10 min of resistance training per day on 5 days a week. ❑ 15 min of walking on 6 days a week.

29. The level of physical activity in patients with chronic obstructive pulmonary disease (COPD) can adequately be reflected by a measurement of:

❑ FEV_1. ❑ Diffusion capacity. ❑ Accelerometer outputs. ❑ Exercise endurance at 75% peak work rate. ❑ 6-min walk test.

30. Which exercise is most fatiguing for the quadriceps muscle in patients with COPD?

❑ Cycling endurance at 75% of peak work rate. ❑ Walking exercise at 75% of peak walking speed. ❑ Walking exercise at 80% of peak walking speed. ❑ Exercise is not fatiguing in patients with COPD.

31. Which exercise test is most responsive to change with bronchodilators in COPD?

❑ FEV_1. ❑ Cycling endurance test at 75% of peak work rate. ❑ 6-min walk test. ❑ Endurance walking test.

32. Which of the following is correct?

❑ Data from exercise assessment can only be used as an exploratory end-point for regulatory authorities. ❑ Data from exercise assessment can only provide mechanistic details to explain the effectiveness of interventions. ❑ European regulators accept exercise assessment as a co-primary end-point in clinical trials. ❑ American regulators accept exercise assessment as a co-primary end-point in clinical trials.

Other titles in the series

ERM 61 – Complex Pleuropulmonary Infections
Gernot Rohde and Dragan Subotic

ERM 60 – The Spectrum of Bronchial Infection
Francesco Blasi and Marc Miravitlles

ERM 59 – COPD and Comorbidity
Klaus F. Rabe, Jadwiga A. Wedzicha and Emiel F.M. Wouters

ERM 58 – Tuberculosis
Christoph Lange and Giovanni Battista Migliori

ERM 57 – Pulmonary Hypertension
M.M. Hoeper and M. Humbert

ERM 56 – Paediatric Asthma
K-H. Carlsen and J. Gerritsen

ERM 55 – New Developments in Mechanical Ventilation
M. Ferrer and P. Pelosi

ERM 54 – Orphan Lung Diseases
J-F. Cordier

ERM 53 – Nosocomial and Ventilator-Associated Pneumonia
A. Torres and S. Ewig

ERM 52 – Bronchiectasis
R.A. Floto and C.S. Haworth

ERM 51 – Difficult-to-Treat Severe Asthma
K.F. Chung, E.H. Bel and S.E. Wenzel

ORDER INFORMATION

Monographs are individually priced.
Visit the European Respiratory Society bookshop
www.ersbookshop.com
For bulk purchases contact the Publications Office directly.
European Respiratory Society Publications Office,
442 Glossop Road, Sheffield, S10 2PX, UK.
Tel: 44 (0)114 267 2860; Fax: 44 (0)114 266 5064; E-mail: sales@ersj.org.uk

EUROPEAN RESPIRATORY *monograph*

The traditional end-points for clinical studies of lung diseases were based on functional parameters. Their value as surrogate markers for disease activity and progression has been increasingly questioned by scientists, carers, regulatory agencies and funding bodies. Novel tools and methods with regard to biomarkers and patient-reported outcomes have made these parameters emerge from their status as interesting secondary end-points and become potential primary outcomes for clinical trials. Nevertheless, their relevance and validity still needs to be proven. This issue of the *European Respiratory Monograph* describes the current status regarding end-points in all relevant areas of pulmonary medicine.

Print ISSN 1025-448x
Online ISSN 2075-6674
Print ISBN 978-1-84984-044-6
Online ISBN 978-1-84984-045-3

Number 62 December 2013
€40.00

ISBN 978-1-84984-044-6